KU Leuven
Humanities and Social Sciences Group
Institute of Philosophy

KU LEUVEN
DOCTORAL SCHOOL
HUMANITIES AND SOCIAL SCIENCES

LIFE'S PLOT

CONTINGENCIES AND SYMMETRIES IN EVOLUTIONARY HISTORY

Hugh DESMOND

Supervisor:
Prof. Andreas DE BLOCK

Co-supervisor:
Prof. Grant RAMSEY

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Philosophy

October 2016

Acknowledgements

What is written last is often what about comes first, and while the fruit of an endeavour invariably receives the most attention, it is the invisible nourishing over the years that is, if anything, more important. It is right that the acknowledgements should preface a book.

First of all, I would like to thank my wife Eleonora, who has kept our new family going during my bouts of absent-mindedness, who has kept on supporting me over the past years, and who – not least – helped substantially with the bibliography work.

Thanks to my mother and father, and two brothers Oisín and William, to whom I could turn for advice, support, help with kids, and – not least – proofreading.

I came to analytic philosophy rather late and almost by accident, and I wish to thank two early mentors for these early stages. First, Paul Moyaert, whom I worked with on Bergson, and from whom I learned to think as concretely as possible. Second, Mark Johnston, for showing how it was possible to do substantial philosophizing in the analytic style, and for showing true companionship in pursuit of wisdom during my year at Princeton.

Once back in Leuven, I knocked on Andreas De Block's door to ask him whether he would supervise work on the question as to whether humans are contingent outcomes of evolution. (I instead eventually settled for the problem of contingency.) I thank him for his patience with my initial stumblings, for taking time to go through many different papers, and for being strict with me on clarity. I learned that clarity of expression is something that must be continually honed.

Two other later mentors I would like to thank. First, Grant Ramsey,

who became a co-supervisor in the later stages of the dissertation, and who has given me valuable advice. Second, Michael Strevens, for penetrating discussions about some aspects of my work when visiting him in New York.

The Higher Institute of Philosophy is a great place to work, with many great people, and I especially wish to thank Jan Heylen, Olivier Lemeire, Markus Eronen, Pieter Thyssen, Stefaan Cuypers, Sylvia Wenmackers, Chris Kelp, Cameron Boult for friendship in philosophy.

Finally, I must thank my good friends Hendrik De Doncker, Takeshi Morisato and Pieter Geeraert for listening to me wax lyrical on occasion.

Overview

Detailed Contents

Introducing Interpretations of Evolutionary History

1 Statement of Problem and Purpose

Just as astronomy teaches us that our planet orbits an unremarkable star at the edge of the Milky Way, evolution teaches us that our species is but a tiny part of the history of life – and, some think, almost an afterthought amidst the vast multitude of species that have ever existed. If we think of the progression of life over time as a complex web of paths, connecting species to species, it would seem easier to find the proverbial needle in a haystack, or drop in the ocean than for evolution to 'find' something resembling the human being in the vast space of biological possibilities.

And yet, find it evolution did. Was this a concatenation of unlikely events, or does life unfold with some directionality over time, biased towards certain types of outcomes rather than others? Once upon a time this question had a straightforward answer: life was like a ladder – the *scala naturae* – with the human as the pinnacle of natural creation. Today, however, the answer is no longer so clear, since we are much more aware of how life could have evolved otherwise.

If evolutionary history were to be replayed from the beginning, what would be the same, and what would be likely different? Would there be a human-like species, multicellularity, or even DNA? This question is the biological equivalent of counterfactual history, such as wondering what would have happened if Caesar had never crossed the Rubicon, or Nazi Germany had won the war. What role does *contingency* play in evolutionary history?

One of the reasons why science alone cannot answer this question for us

is that we have no way of empirically comparing life on Earth with life as it could have possibly evolved. This is what is known as the '$N = 1$ problem' in astrobiology literature: the sample size of all possible evolutionary histories is one. Having a larger sample size, for example, by discovering life on other planets, would allow us to directly test hypotheses about how contingent a given evolutionary outcome is.

However, lack of empirical data is but one obstacle. At a conceptual level, the question – how contingent are evolutionary outcomes – seems so broad and general that one may doubt whether it *can* have any answer. For instance, are some outcomes not much more contingent than others? Any one individual person would seem to be much more contingent as an outcome of evolutionary history (e.g., if your mother and father had never met, you would not be here) than, say, the appearance of multicellular organisms. From this perspective, one may be tempted to say that there is no single answer to contingency in evolution.

Another conceptual problem lies in how we are to understand the evolutionary process itself. Evolution is descent through modification, and such 'modification' can be influenced by any number of causal processes, from mutation to climate change and asteroid impacts. From this perspective, to ask about contingency in evolution 'as a whole' seems to convey only ignorance of the complexity of causal processes that affect evolution.

In contrast to empirical investigation, such conceptual problems lie within the scope and methods of philosophy, and in this dissertation I will seek clarification on conceptual issues that underlie the question about evolutionary contingency. I will pursue the twin goal of analyzing what contingency means, and what the causal basis of evolution implies about contingency.

However, such conceptual problems cannot be pursued wholly independently of empirical phenomena: the analyses of contingency must, at least in principle, be applicable to empirical phenomena. The problem there is that, in order to pursue a full investigation of the question of contingency, I would need to have a full understanding of evolutionary processes – for example, how phylogenetic biases constrain variation and natural selection. However, not only would it be too vast a task to complete a survey of contemporary biological theory, such perfect understanding is not available (for example, it is still controversial as to how precisely phylogenetic

constraints work – and even what those constraints can be).

For this reason, I must limit the scope of the investigation. Instead of directly tackling empirical phenomena, I will focus on what I call *interpretations* of evolutionary history. I will spell out later what precisely I mean by 'interpretation', but for the moment it is sufficient to think of it as a simplified representation of the complex total of evolutionary paths, based on a limited number of principles. By simplifying the complex whole, an interpretation attempts to *explain* some of the large-scale features of evolutionary history, such as whether evolution tends towards any particular kind of state, or towards a limited number of states.

These interpretations are almost always controversial, since they often ignore problematic details for the sake of the coherence of the bigger picture. They are also often very influential, since their conclusions serve to orient how other domains think of evolution. For example, evolutionary meta-ethics has been influenced by interpretations of evolutionary history that imply that outcomes (including moral beliefs) are contingent. Some (e.g. Street 2006) have taken this interpretation to argue that our moral beliefs likely do not correspond to absolute moral truths.

Such interpretations are often products of decades of work by professional biologists grappling with empirical phenomena. Different biologists have built up different interpretations with different conclusions for contingency. Some have argued that evolutionary history tends towards an increase in complexity or size (such as John Bonner or Geerat Vermeij), others (such as Conway Morris) suggest that life inevitably converges on specific structures, and yet others propose that almost everything in life is contingent (such as Gould). These biologists support their claims with many empirical examples. What I will do is to take their interpretations as starting points for my own investigation. In this way, the end-product of scientific investigation becomes the beginning of philosophical reflection.

The reflection undertaken in this dissertation will be structured along two main lines. First, the contingency claims made by biologists are not always justified given their own interpretation of evolutionary history. For example, Gould claims that every single outcome is "utterly" contingent; yet, upon closer analysis, the statements he makes about the course of evolutionary history do not necessarily support this claim. In the first part of the dissertation I will seek to analyze how contingency claims can

be extracted from interpretations of evolutionary history – and how this extraction can sometimes go wrong.

However, I do not wish to take biologists' claims about large-scale features of evolutionary history merely at face value. Investigating the causal basis of interpretations of evolutionary history will be the topic of the second part of the dissertation. In particular, and second, I will be zooming in on what the causal structure of *natural selection* implies about resulting evolutionary histories. I will take a critical look at how some biologists have argued that trends for increased size or complexity result from natural selection, and will propose that they overlook the role that contingency plays in evolution by natural selection. Natural selection does not allow for privileged directions in evolutionary history as a whole, with possibly one exception (to be elaborated later).

This global argumentative structure is summarized in the figure below. In the remainder of this introductory chapter, I will first give a broad overview of interpretations of evolutionary history and then lay out with more precision what I precisely mean by an 'interpretation'. Finally, I will show how the global structure of the dissertation can – at another level – be understood as a sustained response to two skeptical stances towards interpretations of evolutionary history.

2 Some Prominent Interpretations of Evolutionary History

In the following section I will briefly describe four different interpretations of evolutionary history. While each of the four has slightly different implications concerning contingency in evolution, the four form a continuum, ranging from interpreting (almost) every evolutionary outcome as necessary to interpreting (almost) every outcome as contingent. Some of the interpretations are an update of aspects of older interpretations (for example, Bonner's interpretation and the Chain of Being); others (such as the Gouldian interpretation) explicitly react to and distance themselves from other interpretations. After sketching these different interpretations, I will systematize the discussion and give a 'minimal characterization' of interpretations of evolutionary history, which will serve to orient the discussion in the dissertation as a whole.

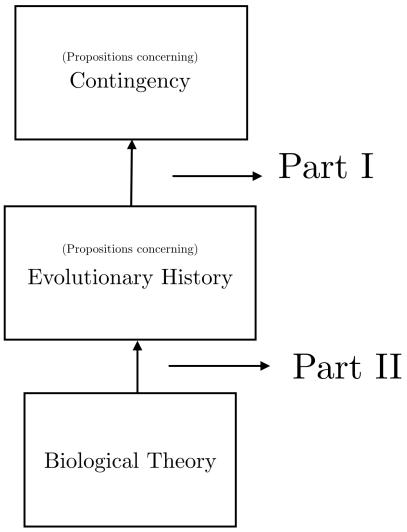

Figure 1: The overall structure of the dissertation, relating biological theory with evolutionary history, and evolutionary history with the contingency of evolutionary outcomes

2.1 The Great Chain of Being

The 'Great Chain of Being' is a term originally used by Arthur Lovejoy (1936) to refer to the metaphysical scheme where all entities are ranked, starting with the inanimate world, rising through plants and animals, through humans and angels, ultimately reaching God. The scheme was popular among 18th century naturalists, such as Charles Bonnet (1720-1793) and Jean-Baptiste Robinet (1735-1820), who used it to interpret evolutionary history as a temporalized *scala naturae*, where the lower rungs of the ladder of life are populated first, followed by higher rungs until the human being evolves (see Ruse 1996: 42-49).

In this way, the Great Chain of Being depended on two now-discarded assumptions: species essentialism, where different species can be identified by a number of essential properties, and a linear progression of nature according to some measure, such as complexity or intelligence[1].

[1]Or simply 'perfection': "one observes a kind of gradation in intelligence of animals, as it exists in the increased perfection of their organization." (Lamarck 1802: 124, quoted

Darwin directly attacked the idea of species essentialism by arguing for the 'slow and gradual' mutation of species: species formed a continuum rather than a discrete set of categories. Species essentialism – including an essentialist understanding of human nature – is now widely seen as incompatible with the phylogenetic definition of species (cf. Mishler and Brandon 1987, Hull 1986). The idea of linear progression was not directly targeted by Darwin, but it became increasingly untenable in the light of his theory of natural selection – and especially after the establishment of the Modern Synthesis in the 1920s and 1930s. Natural selection only distinguishes between organisms that are more or less adapted to a specific environment: it does not care about better or worse in an absolute sense.

It may be noted in passing that not all skepticism towards the Chain of Being idea is merely scientific in nature. First, it has been associated with the ideological use of science – such as creationist science and 'scientific racism' (i.e., the use of science to justify racism). Some, such as David Smith (Smith 2011), even argue that the Chain of Being continues to "cast a long shadow over our contemporary worldview", being "a prerequisite for the notion of dehumanization" (Smith 2011: 42).

Second, some have associated it with a hardwired 'folk biology' – the collection of heuristics and intuitive judgements about living beings (Atran 1998, Gil-White 2001). According to this view, part of the reason we intuitively reach for the Chain of Being metaphor (i.e., understanding humans to be superior to mammals, mammals superior to reptiles, and so on – or even certain groups of humans superior to others) may be because an essentialist appraisal of the value of living beings allows for more immediate judgements on how to act in certain situations (for example, when threatened by potential predators or rival tribes). Thus, even though it is no longer thought useful today, perhaps it was beneficial in our ancestral environment and thus still hardwired into our brain's architecture, making it difficult for us to avoid it.

2.2 The Large-Scale Trends Interpretation

Skepticism towards the Great Chain of Being has not prevented some paleontologists from salvaging one aspect of the interpretation, and making

in Ruse 1996: 50)

it compatible with contemporary phylogeny: the idea that there are global trends in evolution. For example, in his classic work, *The Evolution of Complexity*, John Tyler Bonner argues for a large-scale trend in size and complexity, and even explicitly connects it with the Chain of Being:

> The two generalizations [that we will aim to isolate] are ... (1) an evolution from small to large, and more importantly (2) an evolution from simple to complex. Therefore, our task will be to reexamine this reincarnation, in modern form, of *the great chain of being* from the point of view of current population biology, genetics, developmental biology, and behavior. (Bonner 1988: 25, my emphasis)

For Bonner, selection for size is causally more significant than selection for complexity, because it is the sustained selection for larger body size that drives changes in the organization of an organism. For example, a larger body size means that, at a certain point, the organism cannot only rely on osmosis to take in its nutrients from the environment, but must develop some circulatory system that transports nutrients from the surface to the interior[2].

Further, Bonner hypothesizes that selection for size is nearly ubiquitous: size allows for a competitive advantage in securing nutritive resources, allows for increased evasion of predators, and increased robustness against environmental stresses. Bonner also acknowledges costs associated with increased size (for example, a larger organism needs more nutrition to survive), but argues that, on the whole, selection for increased size dominates.

Bonner uses this hypothesis of the dominance of selection for size to argue that evolutionary history can be understood as a large-scale macroevolutionary trend in size. Increased size drives increased complexity, and drives widespread changes, from life-history to ecology. In this way, Bonner cuts through the complexity of evolutionary history to focus on one parameter – size – and this allows trend to be distinguished from mere

[2]The reason for this is ultimately simple mathematics: volume increases are proportional to the cube of the dimensions of an organism while surface increases are proportional to the square. This relation between body plan and complexity will be an important consideration in chapter 3, when analyzing the Gouldian interpretation.

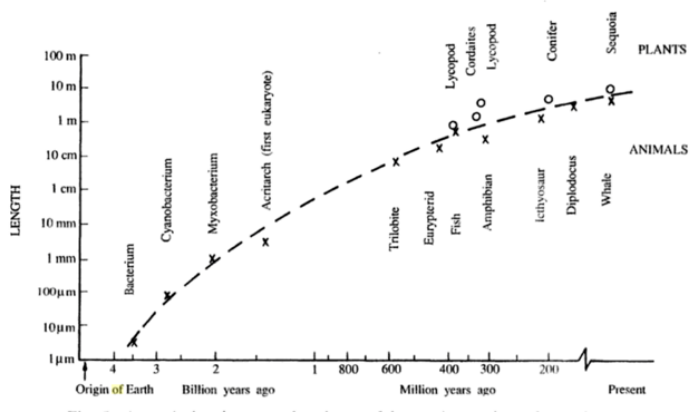

Figure 2: The increase in maximum body size over time. (Reproduced with permission from Bonner 1988.)

fluctuation. The following figure shows how the maximum size of organisms has steadily progressed over time, from the first bacteria to whales and Sequoia trees.

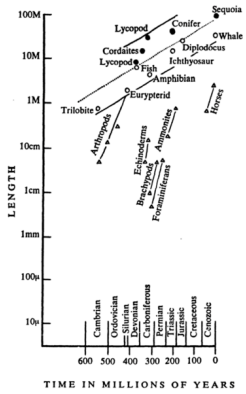

Figure 3: Microtrends in body size, within the large-scale trend in body size. (Reproduced with permission from Bonner 1988).

However, in ascribing such a general trend, Bonner is not denying that there are numerous exceptions to the trend. For instance, if one looks at shorter timescales, one finds that decreases in body size are as frequent as increases, if not more frequent. One main reason for this is that since larger organisms need more energy from the environment for sustenance, they are hit harder during mass extinction events than smaller organisms. This is why, according to Bonner, many of the largest species of families have gone extinct – from club mosses to reptiles – with only the smaller species still extant (Bonner 1988: 39).

Further, such decreases are to be distinguished from the general trend. Sudden increases are also not necessarily part of the trend (see Figure 3). For a microtrend to con-

tribute to the trend in maximum size, the microtrend would need to end at a point where the maximum size increase. Thus, only a minority of microtrends will contribute to the macrotrend.

Although Bonner's case for a trend in size is exceptionally well-documented and argued, numerous other biologists have proposed global macroevolutionary trends. We mention one other prominent account, the Red Queen hypothesis, first formulated by Leigh van Valen in 1973[3]. The Red Queen is a character in Lewis Carroll's *Through the Looking Glass*, who at one point proclaims "Now here, you see, it takes all the running you can do, to keep in the same place." Van Valen, in analyzing the extinction rates of a wide variety of taxa in given adaptive zones, observed that the extinction rate is *constant* and thus independent of the age of the taxon (older taxa are no less likely to go extinct than younger taxa), implying that extinction is a random process. Nonetheless, from an ecological perspective, this does not seem to make sense: if the adaptive zone remains constant over time (i.e. the external environment does not meaningfully change), a population of organisms would be expected to increase its fitness under the influence of natural selection.

Van Valen's solution is to posit that, because of competition and predation, increases in fitness in one lineage come at the expense of the fitness of another lineage. The cheetah increases its fitness by evolving to run faster; correspondingly, the fitness of the gazelle is lowered. This zero-sum game means that the fitnesses of other species will need to increase just so that they will not be driven to extinction. In order to simply maintain a certain ecological status, species will need to continually evolve in order to avoid being at a competitive disadvantage.

The Red Queen hypothesis describes an ecological process, and on its own does not constitute an 'interpretation' of evolutionary history; however, it has been used to this end, most notably by Geerat Vermeij's 1987 view of evolution in terms of 'escalation' – an arms race between predator and prey. The zero-sum game described by the Red Queen has led, according to Vermeij, to a macroevolutionary trend in 'energy-intensiveness': species that extract more resources from the environment and use them in

[3]Incidentally, Valen terms this trend 'A New Evolutionary Law': a further illustration of the affinity between trends and laws.

a more concentrated way (e.g. for predation or competition) have consistently been favoured by natural selection.

Vermeij's interpretation is one of the most worked-out interpretations of evolutionary history, but other commentators, such as Richard Dawkins, have likewise used Van Valen's hypothesis to propose that evolution may legitimately be viewed as directional, despite being driven by 'blind' natural selection (Dawkins 1986).

2.3 The Convergence-Centric View

Another type of trend that natural selection can cause is located at a smaller scale: the trend in a population towards its most adaptive state[4]. When two species independently evolve a similar structure in response to a similar selective pressure, this is known as *convergent evolution*. One classic example of convergent evolution is the similarity in wing design between bats and birds, or the similar morphology of whales, Ichtyosaurs and fish. Another well-analyzed example is the independent evolution of the camera eye in vertebrates and cephalopods. Despite subtle differences between the eyes of, say, octopi and humans (for example, the eyes of octopi do not have a blind spot), the general structure, consisting of pupils, lenses, glassy material and retina, is identical.

Instances of convergent evolution imply that a particular evolutionary outcome, such as the evolution of the camera eye, is not entirely contingent on the past evolutionary path, since two independent evolutionary trajectories could lead to the outcome[5]. The phenomenon of convergent evolution is clearly relevant for questions as to whether a particular outcome is contingent; however, like selection for size or the Red Queen hypothesis, the mechanism of convergent evolution must be further elaborated in order to attain a full interpretation of evolutionary history.

The paleontologist Simon Conway Morris is perhaps the most well-known proponent a 'convergence-centric' interpretation of evolutionary history. In his work he has amassed countless examples of uncanny similarities

[4]Thinking of natural selection as an optimizing process is controversial when one looks in more detail at how natural selection actually occurs in populations. See chapter 5 for further discussion of this.

[5]For more discussion of some problems involved in distinguishing convergent evolution from other types of evolution, see chapter 1.

between lineages which have been evolving independently for hundreds of millions of years. His most spectacular examples include ants that tend to fungus gardens and harvest them at regular intervals, much like human agriculture. He argues that convergence is ubiquitous, and can be found in all the sensory modalities (vision, hearing, olfaction, taste, touch), in cognitive abilities, consciousness and even in culture. In particular, capabilities that seem uniquely human could have been realized by other species. For example, Conway Morris claims that human language "... may, on this planet, be unique, but waiting in the wings of the theatre of consciousness are other minds stirring, poised on the threshold of articulation." (Morris 2003: 253)

In order to argue that a trait that has evolved only once, such as human-like intelligence, is inevitable, or at least highly probable, Conway shows how the *prerequisite* structures for human intelligence have evolved repeatedly, from mimicry and babbling in dolphins to agriculture in ants. Sentience in its various forms, from vision and olfaction to hearing, has evolved repeatedly, as have nervous systems. Human intelligence may only have evolved once, but the building blocks for human intelligence have evolved repeatedly.

We will be analyzing Conway Morris's interpretation in more depth in chapter 3; for the moment we can point to two important parts in his argument. First, he claims by numerous empirical illustrations that convergent evolution is ubiquitous, so much so that "99.9%" of all possible organizations of living beings is maladaptive (Morris 2003: 309). In the words of Dennett (who holds similar views about convergence without elaborating them in the same way as Conway Morris[6]), there are relatively few "Good Moves" in design space (Dennett 1996: 306).

Second, because there are so few structures that actually work, it is 'inevitable' that all of them are realized at some point in evolutionary history. Indeed, many of them are realized multiple times in different lineages. In particular, the prerequisite structures for human intelligence have evolved repeatedly and independently, and Conway Morris takes this to specifically argue that it was inevitable that some species resembling

[6]Dennett interprets natural selection in a more 'standard' way, a way I will associate with the Modern Synthesis in chapter 4. In this view, natural selection can only lead to contingent directionalities.

human beings were to evolve 'sooner or later':

> As all the principal biological properties that characterize humans are convergent, then sooner or later, and we still have a billion years of terrestrial viability in prospect, 'we' as a biological property will emerge. (Morris 2003: 96)

2.4 The Gouldian Interpretation

When it comes to the contingency of human beings within evolutionary history, the contrast between Conway Morris's and Gould's view could not be starker. In Gould's words:

> We came *this close* (put your thumb about a millimeter away from your index finger), thousands and thousands of times, to erasure by the veering of history down another sensible channel. Replay the tape a million times from a Burgess beginning, and I doubt that anything like *Homo sapiens* would ever evolve again. (Gould 1989: 289)

If for Conway Morris evolutionary paths tend to inevitably converge on a limited number of adaptive solutions to design problems, such as wings and limbs, language and culture, and even human intelligence, for Gould, by contrast, nothing is inevitable in evolutionary history. On the contrary, had life on Earth struck out on different paths, especially at the beginning, life might have looked completely different.

Gould made many contributions to biological theory, and other contributions could be mentioned as part of his 'interpretation' of evolutionary history[7]. However, I will isolate two arguments in particular, as these are particularly important for how the contingency of evolutionary outcomes is understood. The first, which I will term the 'disparity reduction' argument, shows how evolutionary history can be irreversibly affected by

[7]In particular, I will not discuss Gould's theory of punctuated equilibrium, as it does not matter for the contingency of outcomes whether speciation events occurred gradually as opposed to discontinuously or very rapidly. We will assume some form of group and species selection in chapter 6, since this is common practice in studies of evolutionary trends, even though these types of selection face some conceptual problems and are not entirely uncontroversial.

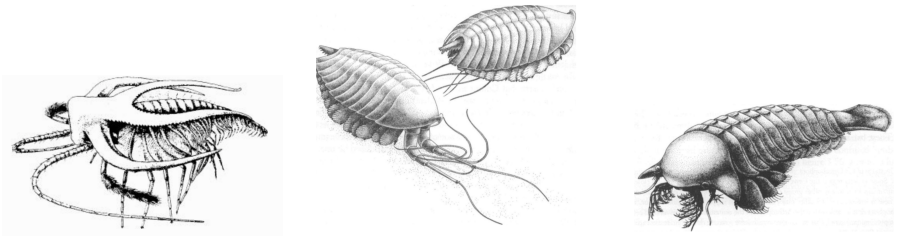

Figure 4: Reconstructions of trilobites. The two left trilobite genuses (Marrella and Leanchoilia) went extinct, and their distinctive body plans with them. The right trilobite (Sanctaris) is thought to be the ancestor of chelicerates (a group including spiders and scorpions). (Reproduced with permission from Gould 1989.)

certain contingent events. The second, here termed the 'random walk' argument, attempts to show how evolutionary trends (in complexity for example) can simply result from biological lineages randomly exploring biological possibilities.

2.4.1 Disparity Reduction

A central observation Gould makes in his 1989 book, *A Wonderful Life*, concerns the proliferation of body plans (*Baupläne*) among the trilobites and other arthropods in the Cambrian. A body plan determines a number of large morphological features, for example how the body of an organism is segmented into different sections, how limbs are placed, and what spatial symmetries the body has. Of the twenty-five basic body plans that can be distinguished, only four remain – a small sample is given in the figure below.
In this way, all *extant* animal life follows just a limited number of body plans, so that the 'disparity' (number of distinct body plans) has decreased in evolutionary history even though the diversity (number of distinct species) has increased.

Crucially, Gould holds that this reduction in disparity occurred during a series of mass extinctions during the Paleozoic, and argues that there is no evidence at all that the body plans that survived were in any way more adaptive (Gould 1989: 236). The selection of organisms during a mass

extinction could be truly random,- resembling a lottery. Alternatively, mass extinctions could represent a radically novel environment, with new selection pressures. In this case, the selection for organisms would be 'effectively' random if viewed from a larger time-scale: since the organisms evolved their traits during normal times, and since the precise nature of the mass extinction is contingent (i.e. could have been otherwise), which organisms are selected for during mass extinctions are contingent as well. There are 'different rules' during mass extinctions (Gould 1989: 306-307).

Only a fraction of body plans survived, but these body plans have *constrained* all future life (the genes, such as the Hox gene, responsible for body plans are very much upstream in development and almost impossible to dislodge through mutation). In this way, the shape of evolutionary history has been largely determined by these 'frozen accidents': a highly impactful but contingent event (a mass extinction) that fixates a number of traits (body plans) that are very resistant to change (due to developmental bias).

Gould allows for convergent evolution, but in the argument above the importance of convergence is undercut. Convergent evolution may have occurred in subsequent evolution, but such convergent evolution only took place in the narrow bands set by the body plans. Had very different body plans been selected at the end of the Cambrian, life today would look very different, regardless of how much or how little convergence there is.

Life is a cone of increasing disparity, but at a certain point this disparity is pared back in a contingent way, in such a way that this contingency is 'passed on' to the evolutionary outcomes. If the tape of life were to be replayed, other body plans could have been selected, and in such a case there would be little to no reason to expect much similarity in evolutionary outcomes.

2.4.2 The Random Walk Argument

Trends in the maximum of some measure (such as complexity) do not have to be caused by any single mechanism. They can also result from a random walk. This is the core of Daniel McShea's distinction between a passive and driven trend. Driven trends show a marked bias towards an increase of complexity during speciation events (represented by the nodes on the graph in Figure 6). By contrast, passive trends arise when the

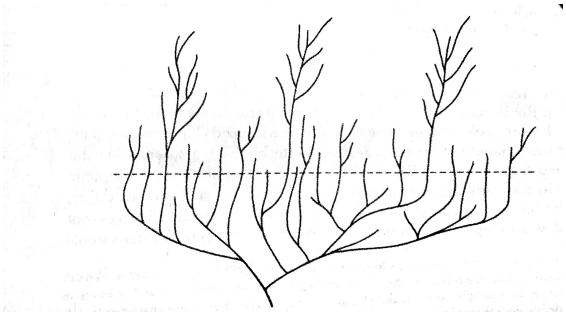

Figure 5: The dotted line represents maximal disparity; after disparity reduction, diversity increases along a select few lineages, but previous level of disparity is never reached. (Reproduced with permission from Gould 1989.)

offspring lineages after an event of speciation will equally tend to decrease or increase in complexity. However, despite this lack of bias, the maximum of complexity will increase regardless, simply by virtue of the expanding variation due to random walk.[8]

Gould uses this distinction to argue that the increase in complexity in evolutionary history is not the result of a selective bias for increased complexity, but simply what Gould calls a 'diffusion from a left wall' of minimum complexity (Figure 7). Another way of putting this is that an increase in maximum complexity (or any trend in some measure) simply could result from a random exploration of evolutionary possibilities.

Besides pointing to how a random walk could have led to the increased complexity we observe, Gould offers two reasons for being skeptical of a driven trend in complexity. First, there is little empirical evidence for a driven trend in complexity (McShea 1996). Second, there is little to no *theoretical reason* to believe complexity should be any more favoured than simplicity by natural selection. In particular, natural selection is an evolutionary mechanism that produces adaptations to particular environments, and in this sense is a local mechanism with no sense of overall directionality. As the environment changes, a previously complex adaptation may

[8]McShea (1994) further distinguishes between strongly driven trends – where the minimum does increase – and weakly driven trends – where it does not. This distinction is not important in this context.

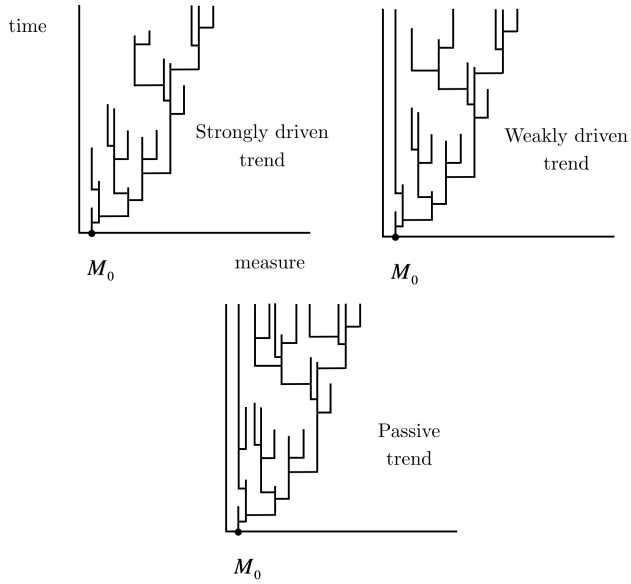

Figure 6: Passive, weakly driven and strongly driven trends.

become useless, and like cave fish having lost the use of their eyes, it may be advantageous that the complex structure loses functionality or even disappears. In fact, Gould holds (1996: 200-1) that there is a good reason to suspect *simplicity* may be generally favoured by natural selection: parasitism is often a successful strategy, and this involves discarding many organ systems (most parts of the digestive system, the sensorimotor system, etc.).

In this way, Gould argues that, even if there are evolutionary trends, such as a trend in complexity, this trend is likely the result of a random exploration of biological space. In this way, Gould is not denying that there are evolutionary trends – he recognizes that they exist – but he is in effect denying that such trends reveal anything *significant* about evolutionary history, for two reasons. First, the trends are a consequence of mathematics, and characterize diffusive processes in general – and so do not reveal anything particularly unique to the causal structure of evolu-

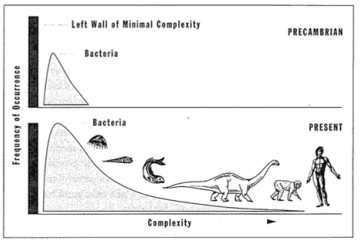

Figure 7: The increase in complexity may be the result of a diffusion from a left wall. (Reproduced with permission from Gould 1996)

tionary history[9]. Second, (following from the first reason) the diffusion from a left wall characterizes any trend in *any* measure – height, weight, pitch of shrieks, number of nervous tics, etc. There is nothing particularly interesting about such a trend in complexity.

2.4.3 Anti-anthropocentrism

An aspect of the Gouldian interpretation of evolutionary history that should be mentioned is the association of trends with a belief in 'progress', namely that more complex organisms are somehow 'better' than less complex ones, and that life thus progresses. In Gould's analysis, this results from an *anthropocentric fixation* on the right tail of complexity. By contrast, if anything, the bacterial 'mode' of life should be deemed the superior: the number of bacterial organisms and species still far outweighs the number of more complex organisms, and on the whole bacteria have been extraordinarily successful in biological evolution, across numerous measures. To consider more complex organisms an 'improvement' is to mistake their marginal importance in life as a whole, simply because we are the ones doing the observation.[10]

[9]This is not to say that mathematical explanations have explanatory value in their own right (see Kitcher 1989, Rice 2011); however, they explain by subsuming a phenomenon under a mathematical pattern, and do not reveal anything interesting about the idiosyncratic causal structure of the phenomenon.

[10]A reader might recognize that Gould here is drawing on what is called observer selection effects, which occur when the patterns one observes in data are correlated with the existence of the observer. This will be discussed later on.

Concerning the evolution of human beings in particular, while Gould acknowledges that a trend in complexity is inevitable, he points out that this does not imply that any specific evolutionary outcome – such as the *Homo sapiens* – should be inevitable. While the existence of right tail is perhaps an inevitable development in evolutionary history, its composition is anything but inevitable: "utterly unpredictable, partly random, and entirely contingent – not at all foreordained by the mechanisms of evolution" (Gould 1996: 174-5). In this way, to conclude the inevitability of humans from the inevitability of increasing complexity is to fall prey to another fallacy born of anthropocentrism: mistaking the small part of the space of biological possibility that we happen to inhabit for a much larger part of that space.

2.4.4 Progress and Distrust of Directionality

The Gouldian interpretation has been very influential, especially among philosophers. Along with its positive claims about the shape of evolutionary history, it is also often seen as grounding a general skepticism towards those who propose directional trends in evolution. Part of this skepticism has to do with the disparity reduction and random walk arguments; however, part of it lies in how biologists who propose some directional trend are distrusted as harbouring a 'surreptitious' belief in evolutionary progress. A good definition of progress is that it is a directional change plus an evaluative judgement about that change, one life form being 'higher' or 'better' than another (Ayala 1974). The Gouldian distrust then leads to a suspicion of biologists who propose directional trends are motivated by evaluative judgements.

The most elaborate development of this scheme is Michael Ruse's *From Monad to Man*, where Ruse traces how the concept of progress has informed biologists, up to the present day:

> Progress broadly understood continues to shape the thinking of some of the most professional of evolutionary biologists. The mature science has cultural undertones. (Ruse 1996: 484)

So while the claim that the idea of progress influenced *some* 19th and 20th century biologists is uncontroversial, Ruse goes further and believes one can recognize it in some contemporary work as well.

Some biologists, such as Normal Rockwell and Anthony Hammam, overtly believe in some form of anthropocentric progressionism. For example, consider Hammam in a 1991 interview:

> I think that man is definitely a success story - but you've got to define your criteria. Consciousness and the ability to analyze the rest of the world and ourselves in the way that we do. The gift of language - all the usual things - is amazing. In those senses, yes, we certainly can be regarded as a pinnacle. (Hallam 1991, quoted in Ruse 1996: 493-4)

While such statements are in need of little interpretative imagination, Ruse makes a leap to find similar implicit assumptions ('sliding' towards progress) in the work of Van Valen and Vermeij. Ruse finds "hints" of "slides from the comparative to the absolute" in the following passage by Vermeij:

> It is possible, however, that species have improved in their capacity to survive in the *physical* environment. Many of the characteristics associated with competitive and defensive superiority-large body size, high body temperature, parental care of the young, and a tightly sealing exo-skeleton, for example-also buffer individuals against short-term fluctuations in temperature and other physical factors (Vermeij 1987: 421, quoted in Ruse 1996: 493).

The implication is clear: biologists, such as Vermeij, while highly trained and knowledgeable, are *still* subtly informed by primitive and outdated notions of progress. It is similar to the functioning of implicit racist or sexist biases: even though we may not be consciously aware of these biases, our judgements are still influenced by them.

Commentators such as Ruse seize on such research as evidence of an ongoing anthropocentric bias in the interpretation of patterns in macroevolution. Despite the considerable progress in evolutionary biology, Ruse writes *'Plus ça change, plus c'est la même chose'* (Ruse 1996: 494): despite extensive scientific training, even professional scientists hold fast to the idea of progress, and sometimes even to some form of anthropocentrism.

However, there is an argument to be made that Ruse misinterprets Vermeij. After the quoted passage, Vermeij continues:

> Consequently, individuals are able to carry on normal activity, or at least to survive, when conditions are temporarily unfavorable. Without such characteristics, individuals would be able to persist in a much smaller range of physical conditions. (Vermeij 1987: 421)

While Vermeij's general analysis can be argued to be flawed in other ways (see chapter 6), it is safe to say that his point does not concern increases in lifespan (or other changes in life-history structure, for that matter). Instead, Vermeij points out that there may be a trend in adaptations that allow an organism to be viable in an *increasing range of physical environments*. Vermeij's argument is considerably more nuanced than simply introducing an absolute measure of progress; instead, he is working entirely within the framework of the local nature of natural selection, and proposes only a trend in adaptations that *expand* this locality. There is no jump towards the absolute, merely an expansion of the comparative. Thus it is at least uncharitable and even arguably wrong to accuse Vermeij of 'sliding' from comparative to an absolute sense of progress.

Further, in Ruse's eagerness to see a 'slide' to Progress whenever an evolutionary biologist proposes an evolutionary trend, Ruse himself is committing a 'slide' towards the belief that the scientist must be informed by some cultural belief and must be therefore biased. In a review of Ruse's book, Hull notes that "when he [Ruse] turns to modern biologists, he weaves his conceptual net so finely that almost no one can avoid being classed as a progressivist" (Hull 1997:497).

Such a fine net can also be harmful, if, to use another metaphor, it leads to the baby being thrown out with the bathwater. Theoretical and empirical investigation into large-scale macroevolutionary patterns is valuable scientific work, both because it deepens our understanding of evolution as a whole and because of its philosophical consequences. From this perspective, an oversubscription to the Gouldian interpretation may actually stifle scientific investigation by placing a taboo on certain research[11].

[11] Ironically, Gould himself has been accused of smuggling cultural values into his scientific pronouncements. Conway Morris pointed to certain libertarian beliefs influencing Gould's views on contingency:

> In brief, his assessment of Man as an evolutionary accident is to lead us

3 Minimal Characterization of an Interpretation of Evolutionary History

If interpretations are simplified representations of evolutionary history, the preceding discussion suggests they are simultaneously quite multifaceted: influenced by ideological presuppositions and yet grounded in biological theory, explanatory of large-scale features in evolutionary history and yet subject to many exceptions. In this section we will give a minimal characterization of the most important elements involved in such interpretations, and we will use this characterization to unify some of their various properties, most notably their peculiar explanatory nature.

There are some precedents to the concept of an 'interpretation of evolutionary history'. Sterelny, when discussing the clash between Gould and Dawkins, refers to 'perspectives' on life (Sterelny 2001). Powell has referred to the 'Gouldian view of life' when discussing Gould's views on evolutionary contingency (Powell 2009, 2012). However, it is unclear to what extent these authors intend 'perspectives' or 'views' merely as metaphors, for they never precisely stipulate what constitutes a perspective or view. For example, Gould and Dawkins have different opinions about evolutionary contingency, but in what way precisely is this difference in opinion simply a matter of opinion? There is also no sustained engagement with obvious skeptical stances towards whether a unified perspective or view of life is even *possible*.

As already explained, the working assumption of this dissertation is that these are not mere metaphors, and constitute legitimate efforts to

---

> into a libertarian attitude whereby, by virtue of a cosmic accident, we, and we alone, have no choice but to take responsibility for our own destiny and mould it to our desire. (Morris 1998: 14).

For the sake of completeness, we should also mention that Conway Morris has been accused of smuggling in religious beliefs. For example, in a review of Conway Morris's book *Life's Solution: Inevitable Humans in a Lonely Universe*, Donald R. Prothero writes:

> One wonders if he [Conway Morris] is writing for the judges of the Templeton Prize, which is awarded to those who try to reconcile science and religion, rather than for an audience of people interested solely in scientific questions. (Prothero 1997: 57)

interpret biological theory in order to arrive at a general conclusion about the contingency of evolutionary outcomes. The two parts of the dissertation will be a sustained testing of this working assumption, but in order to begin, I need to stipulate a characterization that at least is *prima facie* plausible and that can be used for the purposes of the present discussion.

Note that what seems like a *prima facie* plausible definition proves almost never to be so upon further reflection and analysis, and doubtless this will also be the case here. However, I do not wish to engage in a philosophical investigation of the type 'What is X', with X being an interpretation of evolutionary history. I wish only to capture the *main features* of the interpretations of evolutionary history described in the preceding section. Perhaps this minimal characterization does not cover all particular instances, but this is a line of investigation that is not necessary for the purposes of the dissertation. I wish merely to use the minimal characterization as a basis for reflection on how interpretations relate to contingency, and how evolutionary dynamics relate to interpretations.

In the view I will take, interpretations of evolutionary history attempt to specify some set of evolutionary trajectories that is somehow *representative* of all possible evolutionary histories. Actual evolutionary history is a complex and intractable mass of evolutionary trajectories; an interpretation of evolutionary history cuts through this complexity and selects only a few trajectories which are then claimed to be present in most if not all possible evolutionary histories.

It is this simplification of actual history, and placement within a space of possible evolutionary histories that allows for contingency claims to be made. For example, if a convergence of evolutionary trajectories is deemed representative for all possible evolutionary histories, then the outcome of that convergence is not contingent. Without comparing the actual to the possible, claims about the contingency of an actual (or possible) evolutionary outcome cannot be made.

As we will see in chapter 3, not all biologists are equally explicit about this aspect of their interpretation. Conway Morris is more explicit about the nature of possible evolutionary histories than Gould is, and the Gouldian interpretation needs more reconstruction. However, this contrastive aspect of an interpretation – contrasting actual evolutionary trajectories with other trajectories that could have been taken – is crucial, and forms

the basis for claims about the contingency of individual outcomes.

Metaphors of life's history often reflect interpretations. Most obviously, the *ladder* of life reflects an interpretation of evolutionary trajectories either as converging on a particular outcome (the human being), or as characterized by a strict directionality (for example, pointing towards increased complexity or intelligence). However, also the more recent metaphor, the *tree* of life, captures how evolutionary history typically diverges with a steadily increasing expansion in variation between species. Yet the metaphor of the tree has been criticized as still too directional and underestimating contingency, since it does not capture the number of evolutionary dead-ends there have been. Hence the metaphor of the *bush* was a short-lived alternative to the tree (most notably touted by Gould), given that it captured phenomena such as disparity reduction. Finally, in light of how interlinked evolutionary trajectories are, the most common metaphor today is that of a *web*.

In this way, interpretations of evolutionary history imply a certain shape or topology of evolutionary trajectories: they tell us what direction, if any, evolutionary trajectories tend to take in possibility space; they tell us how paths cluster together and how they spread out over time. Ladders, trees, bushes and webs are all different ways of synthesizing a complex topology into a single image.

At a finer level of analysis, the evolutionary trajectories are determined by two elements: a *possibility space* (which at least defines what the initial states are and what the outcome states are) and *causal transitions* between states in that possibility space. Possibility space and transitions together define a causal network, and in this way the representative set of evolutionary trajectories can be thought of as a *representative causal network*.

The transitions in evolutionary history are all relations of descent with modification; however, interpretations of evolutionary history make claims about what the key mechanisms are that drive these transitions. For Conway Morris and Bonner, the key mechanism is natural selection; for Gould, there is an absence of any driving mechanism (unless one wishes to count drift as a mechanism[12]; this is not what Gould intends when using Mc-

[12]We discuss the relation of selection to drift in chapter 6. Note that I use the term 'mechanism' loosely (following common biological practice), and not as a technical

Shea's concept of a passive trend), but developmental biases do play an important role in constraining transitions. Other mechanisms (some of them now discredited) have been included in other contemporary and historical interpretations of evolutionary history that we have not here discussed: mutational bias (Louis 2016), orthogenesis (Eimer 1898), creative force (Bergson 1911), self-organization (Kauffman 1993), entropy-increase (Brooks and Wiley 1988).

It is in this respect that interpretations of evolutionary history are not simply typical 'narrative explanations' of history. A narrative explanation, as it is often understood, explains a particular outcome in terms of a path leading to that outcome (i.e., in terms of the outcomes preceding states), sometimes contrasting the particular path to other paths that could have been taken. In specifying a representative causal network, interpretations resemble narrative explanations.

A narrative explanation is often contrasted to law-based explanations, which explain an outcome in terms of a universal law and the conditions of application – thus allowing for prediction. Narrative explanations do not allow for prediction. In specifying a limited number of key mechanisms that cause the shape of the representative causal network, interpretations allow for a certain degree of prediction. For example, an interpretation such as that of Conway Morris would make certain predictions about how life on alien planets would look. Bonner's interpretation would make predictions that possible evolutionary histories would be characterized by a trend in size.

Even the Gouldian interpretation, by emphasizing the 'utter' contingency of evolutionary outcomes, would predict that life on alien planets would look radically different to life on our planet. In fact, an observation of similarities in outcomes despite difference in evolutionary trajectories would count strongly against the Gouldian interpretation.

A metaphor for this aspect of interpretations of evolutionary history, deepening the connection between laws and narrative, would be the *plot* of a story. A good story has properties that are remarkably similar to good explanation. This old insight goes back to Aristotle, who in his *Poetics*

philosophical term. I do not wish to engage in the current debate whether or not drift and selection can be understood as mechanisms (e.g. Skipper and Millstein 2005, Barros 2008).

argued that a good plot shows how the crucial events in the story "must come from the structure of the plot itself, so that, from what has happened before it turns out that these things would necessarily or probably happen" (Aristotle 1997:87). Interpretations of evolutionary history, in searching for the key mechanisms that hold together important features in the shape of life's history, can in this regard be seen as searching for the 'plot' of 'life's story'.[13]

Summary

From this perspective, 'interpretations of evolutionary history' can be defined by the following three elements:

(1) A *possibility space*, which determines how evolutionary states are to be characterized – most importantly, the initial states and the outcome states.

(2) A hypothesis about what *key mechanisms* determine the transitions in possibility space.

(3) A *representative causal network*, consisting of a set of possible evolutionary trajectories connecting up the different states in possibility space.

These three elements are not necessarily sufficient for characterizing a global view; particular interpretations may be characterized by additional elements. Hence this proposal may be understood as a *minimal characterization* of global views of evolutionary history. In particular, (1) and (2) are grounded in biological theory; and questions concerning contingency can be analyzed solely as structures inherent in the representative causal network.

[13]Not all good stories – especially Modernistic ones – have a plot of course, and hence contemporary narrative theorists loosen the requirement of a tight plot, and require only 'thematic unity' of a good narrative (Currie 2010). This would probably be a more accurate description of the function that the key mechanisms play in interpretations of evolutionary history, but we are more interested in the metaphor, and the metaphor should not be driven too far either. After all, here is little to no equivalent of the predictive aspect of interpretations in storytelling.

Further, these elements correspond to different aspects of the peculiar explanatory character of an interpretation of evolutionary history (including how they can be influenced by ideological concerns):

(i) Interpretations have some predictive value: they extrapolate to possible evolutionary histories, based on element (2), the hypothesis of what the key causal mechanisms are.

(ii) Interpretations are also narratives: they explain an outcome by tracing a path that a system (an individual lineage, or evolution as a whole) has followed.

(iii) Interpretations can be influenced by *ideological presuppositions*: because the reasoning process leading to (3) is often blackboxed, it can be influenced by religious/political beliefs.

4 Two Skeptical Stances

If we consider how prominent scientists, with access to the same theoretical knowledge and to the same empirical data sets, can arrive at diametrically opposed views of evolutionary history, and never resolve this conflict, a skeptical stance towards the plausibility of such global views becomes tempting. According to such a skeptical view, Gould and Conway Morris engage in unfalsifiable speculation and hence it is not surprising that this leads to an irresolvable conflict. In reality, it is not possible to have a coherent 'global view' of evolutionary history that brings together our understanding of evolutionary states and mechanisms in order to inform claims about evolutionary contingency. Perhaps claims about the contingency of *particular* outcomes could be legitimately made, but definitely not claims about the contingency of evolutionary outcomes in *general*.

There are two main ways of filling out this general skeptical response, and each will form a major issue that we will engage in this dissertation. The first skeptical response would be to point to conceptual confusion at the heart of the disagreement. As we will see, Conway Morris is interested only in the occurrence of certain properties in evolutionary history, such as perception, locomotion, intelligence and culture, while Gould is interested in the occurrence of species in their full detail. Put differently,

Conway Morris is interested in constructing some kind of 'periodic table' of convergent structures, where the *Homo sapiens* might be one particular representative of the slot 'intelligent, cultural being'. Gould is interested in *Homo sapiens* in its full detail, and hence it is not surprising that Gould concludes that the occurrence of the human being is more contingent than does Conway Morris.

This skeptical response points to the *description-dependence* of contingency: contingency comes in degrees, and you get different answers according to how you describe the process. There is description-dependence of the outcome, where a fine-grained description seems to go hand in hand with greater contingency, but there is also description dependence of the initial conditions. How far back in time do we have to go to judge the contingency of human beings? Back to the last common ancestor with the chimpanzees, or back to the trilobites of the Cambrian, or back to the very first living being? And once we have decided on the general type of initial condition, how precisely do we specify the initial condition? Must we only specify the phenotype of the initial species, or also the genotype, developmental processes and environmental states?

In this way, the skeptic may simply say that questions about evolutionary history 'as a whole' are simply hopelessly vague. They do not have a single answer: it depends how you describe the evolutionary outcomes and the initial conditions. Without specifying precisely that to which we are referring , irresolvable clashes such as that between Gould and Conway Morris are not surprising. They use terms with ambiguous reference, and hence the conflicts might well be pseudo-conflicts that can be dissolved through conceptual clarification.

A second, related, way of filling out a general skeptical response would be to show how general claims about evolutionary history always turn out to be false. This response piggy-backs on an influential view in the philosophy of biology, namely that there are no universal generalizations in biology, only contingent generalizations: either a generalization is based on mathematics or laws of physics and is universally valid, or it depends on the existence of some contingent entity originating in evolutionary history (Beatty 1995). For example, even Mendelian genetics, with its elegant rules, depends on the origin of the DNA molecule and thus would not necessarily reveal anything about an evolutionary history where DNA had

never emerged. At best, one can obtain generalizations that are *locally* true, i.e., true for particular entities given particular conditions. However, for evolutionary histories where these conditions do not obtain, or these entities do not exist, such generalizations are inapplicable.

Hence arises a skepticism towards generalizations about *all possible* evolutionary histories (excepting generalizations that are based on mathematics or the laws of physics). A global view generalizes about large-scale tendencies in evolution, and uses such generalization for further general claims about the contingency of evolutionary outcomes. Such generalizations themselves are necessarily contingent: limited to parts of actual evolutionary history, but not necessarily open to extrapolation to all of evolutionary history or to possible evolutionary histories.

This skeptical response is related to the first response in that the lack of reference of generalizations can be avoided by reducing the scope of the claim. General claims about evolutionary history as a whole are independent of initial conditions, and all such claims turn out to be false. Hence, by introducing a specific characterization of initial conditions (and of the outcomes), we can obtain true claims about *parts* of *some* evolutionary histories.

5 Structure of the Dissertation

The overall purpose can be reframed as a sustained engagement with these skeptical responses, in order to obtain more robust conceptual foundations for such global views of evolutionary history. Dealing with the first skeptical response, I will be asking how claims about contingency in evolutionary history change as the phenomena are described differently. Dealing with the second response, I will be investigating the relation between natural selection – the mechanism which many prominent global views draw on – and contingency. To what extent does natural selection give rise to directional evolution, and to what extent are such trends contingent on environmental circumstances?

It is important to emphasize that the goal of the dissertation is *not* to provide an interpretation of evolutionary history. Such an endeavour is arguably not the task of a philosopher, but of a practising biologist, who would be able to flesh out a conceptual framework with a deep knowledge

of empirical details. Instead, by analyzing fundamental concepts such as contingency and natural selection, the goal is to outline the boundaries of possible interpretations of evolutionary history that are not vulnerable to the skeptical responses outlined above.

This project fills something of a lacuna in the philosophy of biology literature. As mentioned, authors have alluded to perspectives and views of evolutionary history, but there is currently no dedicated analysis of the conceptual foundations of global views of evolutionary history, and how they do not necessarily succumb to the skeptical responses described above.

While the work hopes to break new ground, it will be drawing heavily on a number of directly relevant literature. Of prime importance is the literature dedicated to the analysis of the concept of contingency in macroevolution (e.g. Beatty 2006, Beatty and Desjardins 2009, Desjardins 2012), and in dialogue with this literature, I will be offering a novel analysis of the concept of contingency, in terms of symmetries and breakings of symmetries in causal networks. However, in the context of the dissertation, the ultimate purpose of this analysis will not be just to achieve clarity on the meaning of contingency for its own sake, but to map the different ways that claims about contingency can be drawn from global views of evolutionary history. Along the way I hope to point to some common errors in the way this is done.

In other words, I will not be analyzing what role contingency 'objectively' plays in empirical evolution. This would entail a detailed analysis of the latest theories of developmental processes, and their relation to natural selection and to gene networks. This question is, arguably, intractable in its full generality, and in any case is a task for a biologist rather than a philosopher. In this way we need not be primarily concerned with related debates, such as issues in experimental evolution (such as Richard Lenski's Long Term Evolution Experiment), or certain aspects of the adaptationism debate, such as the question whether natural selection can overpower developmental constraints or not.

Instead, my 'raw data' will be the interpretations of evolutionary history that various biologists, such as Bonner, Gould or Conway Morris, have arrived at through long and pain-staking grappling with the empirical phenomena. I will take their views for granted as starting points,

and, through a renewed clarification of what contingency means, examine what their views imply about the contingency of evolutionary outcomes. This is a worthwhile inquiry, since biologists have sometimes overstated conclusions about contingency, or have confused contingency with related concepts such as probability.

The second main body of literature I will be drawing upon is that on large-scale trends in evolution. There is a long-standing tradition in evolutionary biology that searches for tendencies that hold universally across all possible evolutionary histories, tendencies that sometimes are even termed 'laws' of evolution (such as the Red Queen Hypothesis by Valen 1973). Such universal tendencies do not necessarily need to be driven by natural selection (e.g. McShea and Brandon 2010), but natural selection has played the key role in many of the most prominent proposals: trends towards increased size, and towards increased energy-intensiveness (e.g. Vermeij 1987, Bonner 1988). Bonner and Vermeij take large-scale trends driven by natural selection as a basis for a view of evolutionary history, and draw conclusions concerning the contingency of evolutionary outcomes.

However, there has also been an equally long-standing skepticism towards such trends (Dawkins 1986, Ruse 1996, Gould 1996), and I focus on a particular aspect of this skepticism concerning whether the outcomes of evolution by natural selection can be anything other than contingent. In re-analyzing the relation between natural selection and contingency, I will take into consideration recent advances in the Extended Synthesis, which highlight previously ignored aspects of the relationship between environment and phenotype (plasticity and niche construction).

Analogous to the first part of the dissertation, the purpose in the second part will not be to propose a rival candidate to the large-scale trends in size, energy-intensiveness or complexity. Rather, taking theories about the causal structure of natural selection as our 'raw data', I will be inquiring as to what conclusions about the contingency of outcomes of natural selection *could* be drawn. In particular, I will outline the possibility that the tendency towards increased plasticity is non-contingent – in the sense that it is common to all *possible* evolutionary histories – but this will not constitute a claim about *actual* evolutionary history, where the tendency could be counteracted by other evolutionary mechanisms.

Part I

Analyzing Contingency Claims

Chapter 1

The Challenge of Description Dependence

In order to answer the skeptical challenge of description dependence in particular, the first part will adopt the following argumentative structure. In the current chapter I explain the skeptical challenge in more detail by inserting it both into a broader philosophical perspective as well as into the specific debate about the role of path-dependence in explanations. As will become clear, this debate provides the ideal context for a systematic analysis of claims about evolutionary contingency.

The second chapter aims at providing just this analysis. In particular, I will propose that contingency claims can be analyzed in terms of symmetries in causal networks. While this is a novel analysis that solves some outstanding problems within the specialized literature on path-dependence, the analysis will also provide the ideal tools with which to analyze interpretations of evolutionary history. The third chapter will attempt to show how this can be done with proper precision. To that end, I analyze the Gouldian and convergence-centric interpretation in terms of causal networks, with special attention to how it can help clarify some important ambiguities inherent in those interpretations.

This is the argumentative arc followed by the whole first part. Focusing on the structure of the current chapter, we will begin in the first section by situating the challenge of description-dependence by zooming out to the broadest (though yet relevant) philosophical horizon: the issue of Hegelian-

type interpretations of history. Such interpretations bear significant resemblances to interpretations of evolutionary history, and it will be instructive to connect the skeptical stances (not only description-dependence, but also causal complexity) with skepticism towards Hegelian-type philosophies of history. In particular, we will discuss Danto's skepticism towards what he calls 'substantive' philosophies of history.

In the second section, we will continue with Danto's account of 'narrative explanations' as his preferred account of explanations of history, and contrast it with Hempel's and Oppenheim's law-based model of scientific explanation. Both Danto and Hempel thought that narrative explanations as such were not scientific (in Hempel's perspective, narrative explanations were simply incomplete scientific explanations). If philosophy of science had not evolved from its status in the 1960s, it would have been difficult to give a thoroughgoing analysis of how biologists interpret evolutionary history, since the type of explanations they look for are neither Danto's narrative explanation, nor Hempel's law-based ones.

However, developments subsequent to the 1960s in philosophy of science have cast considerable doubt on this dualism between what I call nomothetic and narrative poles of explanation. In particular, the causal approach can be seen as integrating some elements that are typical of narrative explanation. This causal approach is also where path-dependent explanations are to be situated: path-dependent explanations are, in some respects, fully elaborated causal explanations. In this way, interpretations of evolutionary history can integrate both narrative and law-based elements (as outlined in the introduction), and yet escape Danto's dismissal of such combinations as an "intellectual monster" (Danto 1965:15).

The third section will focus on the diverse manifestations of path-dependence, in order to develop intuitions concerning what precisely path-dependence is (and how it relates to contingency). Then, in a final section, we will formulate the challenge of description-dependence as a three-fold challenge that any account of path-dependent explanation must meet.

1 Background: Skepticism towards Substantive Philosophy of History

Disciplines such as paleontology and anthropology, but also political science and sociology do not seek grand unifying laws, but instead try to reconstruct past sequences of events. An important element in this reconstruction involves categorizing past events. A specific example of this (cf. Falk 2011): when Raymond Dart found a fossil skull in 1925, he needed to decide whether to categorize the fossil as belonging to an ancestor of modern humans, or to an independent lineage of ape. Based on the evidence (e.g. the morphological properties of the skull), he made the judgement to categorize it as an ancestor of modern humans, and named this genus *Australopithecus*. In this way, Dart made a different reconstruction of past events (a different 'narrative explanation') than his critics, who, by contrast, categorized the fossil as belonging to an ape.

In general, evolutionary biology is, to a significant degree, a *historical* discipline, in the sense that it, unlike physics, is not concerned with eternal and universal laws, but with actual sequences of events. This historical character of science was recognized early on by philosophers of science (in the 1940s through 1960s), but it was seen as a problem, primarily because narrative sequences, in contrast to laws, do not allow for *predictability*. This predictability, especially for the post-logical empiricist philosophers of science such as Hempel, was seen as a sine qua non for scientific explanation (distinguishing science from pseudoscience), and much effort during this time was expended in trying to analyse the 'logic' of narrative explanation (e.g. Hempel 1942, Gardiner 1952, Scriven 1959).

However, interpretations of evolutionary history go further than simply reconstructing actual sequences of events. They make additional judgements as to how possible sequences of events converge on similar end points (Conway Morris), as to how possible sequences tend to evolve along certain directions (Bonner) – or as to how possible sequences do not either converge or tend in certain directions (Gould). And, as mentioned in the introduction, they have predictive implications, for example concerning what general properties alien life might be expected to have. This aspect of interpretations of evolutionary history, provokes, if anything, even more skepticism, not least because it has Hegelian undertones.

1.1 The Hegelian Shadow

There are some striking similarities between the Hegelian approach to history and interpretations of evolutionary history. For Hegel, history unfolds according to the principle of *dialectics*, which refers to a pattern consisting of three interrelated stages of a system: an initial state where the parts of a system are relatively uniform and undifferentiated, an intermediary state characterized by tensions and oppositions between the various parts, and a more holistic final state where previous oppositions are integrated into a complex whole consisting of differentiated parts.

How precisely this dialectical principle should be interpreted need not concern us here[1]; what is important is that dialectics characterizes the important transitions in history, but not necessarily all transitions. As Danto notes, "Nothing that happened in Siberia, for instance, was considered by Hegel to be part of history" (Danto 1965: 14). In the parlance of this dissertation, this is reminiscent of how an interpretation of evolutionary history conveys a *representative* causal network, which does not necessarily correspond to the totality of evolutionary trajectories in actual evolutionary history. The dynamics of dialectics dictates which transitions in history are to be ignored, and which can be integrated, in order to explain the overall shape of history (which for Hegel entails a directionality towards self-actualization and freedom, but what precisely the shape of history is for Hegel is of less interest to us than his general approach to history).

Hegel made a number of controversial assumptions, for example, his identification of the dialectics of history and the dialectics of thought. Nonetheless, as a general approach to interpreting history, Hegel had a profound impact on subsequent thinkers. For instance, Marx and his followers in the 20th century adopted dialectics as a principle of the dynamics of history while at the same time rejecting its idealistic nature (instead reconceiving it as class struggle).

[1] It is not a 'causal mechanism' in the way natural selection is, but rather some kind of abstract pattern that is instantiated across diverse phenomena. Another complicating factor is that it is not only a pattern that characterized the dynamics of history, but also the dynamics of thought. This identification of the dynamics of history with the dynamics of thought is an important factor underlying his dictum 'the Real is the Rational' (Aboulafia 1980, Hegel 1996).

By contrast, in early analytic philosophy, history became a relatively marginalized topic, partially in reaction to Hegel. The Hegelian philosophy of history was seen as encroaching on lines of investigation that belonged to other disciplines and was seen as excessively speculative. The analytic philosophy of history had a period of flourishing in the 1950s and 1960s, and it is instructive to consider the criticisms of one prominent exponent, Arthur Danto.

1.2 Danto's Skepticism

The first thing Danto does, in his central work *Narration and Knowledge* (1964), is to explicitly distance his approach to the explanation of history from what he calls a *substantive* philosophy of history. According to Danto, a substantive philosophy of history – represented by exponents such as Hegel or Marx – aims to interpret the *whole of history*, in such a way that future events can be *predicted*. Such a philosophy of history, Danto writes, is

> An intellectual monster, (...) which is neither history nor science, though it resembles the one and makes claims for itself which only the other can make. (Danto 1965: 15)

As we will see, for Danto, a substantive philosophy of history is a category mistake: it attempts to achieve the historian's perspective on the *present*, and this is impossible. The historian's perspective is only epistemically accessible when the events lie in the past; an *actor* in history has no way of knowing the 'historical significance' of present events, since the events that make the present 'historically significant' may still lie in the future.

If we dig deeper, we can identify two key motivations underlying Danto's claims, each of which bears an interesting resemblance to the two skeptical stances we have identified in the introduction. First, the issue of description-dependence:

> There is an unexpungeable factor of *convention* and of *arbitrariness* in historical description, and this makes it exceedingly difficult, if not impossible, to speak, as the substantive

> philosopher of history wishes to, of *the* story of the whole of history, or, for that matter, *the* story of any set of events. (Danto 1965: 15, first two emphases mine)

And, second, the challenge of causal complexity – manifested through a failure to predict: "Existing philosophies of history are unspeakably inept, with almost no power to predict." (Danto 1965: 5) Danto supports this conclusion with some kind of pessimistic meta-induction: perhaps actors in history, armed with a powerful substantive philosophy of history, might be able to predict which of the current events will turn out to be significant on a historical scale. However, all substantive philosophies of history in the past have failed to provide this kind of predictive power.

By contrast, an *analytic* philosophy of history – the type of philosophy of history that Danto wishes to outline in his account of narrative explanation – is not concerned with the subject matter of history itself, but rather with the logical structure of the claims made by historians.

1.3 Hempel and Danto: Nomothetic and Narrative Explanation

In Danto's view, a historian setting out to explain the occurrence of the First World War will not be content with theories obtained from anthropology and sociology about what antecedent conditions tend to encourage warfare between groups of humans. The historian will want to explain the First World War in its *idiosyncratic detail*.

Yet, the historian will not simply want to list all possible details. Danto imagines the 'Ideal Chronicle' of the whole past to be an enumeration of all past facts in exhaustive detail. This Ideal Chronicle is like some Humean mosaic: a collection of particular facts. For example, the Ideal Chronicle of the French Revolution would enumerate the thoughts and deeds of every single person involved in the event.

What then is involved in historical explanation that is missing from simply describing the Ideal Chronicle? According to Danto, it is the viewpoint of the historian, of a person who looks back onto historical events. The Ideal Chronicle may convey the thoughts and deeds of every single person at the time, but what a historical explanation attempts to target

is not how some historical event was experienced at the time, but rather how its significance is to be gauged in light of subsequent events.

This retrospective nature of historical explanation is closely related to the non-verifiability of 'narrative sentences'. For example, a historian might claim 'Petrarch climbing the Mt. Ventoux was the start of the Renaissance'. Against Ayer, who believed that only verifiable propositions are meaningful, Danto argued that, even though this proposition is meaningful to historians (it is not nonsense), it is not verifiable. A contemporary might have gone out and observe whether Petrarch was in fact climbing the Mt. Ventoux, but they would not have been able to observe whether this act was the start of the Renaissance.

The retrospective character of narrative sentences means they do not refer to events localized in time, but instead to many events extended over time. Narrative sentences are not listed in the set of propositions contained in the Ideal Chronicle. Another example Danto frequently uses is the proposition 'The Thirty Years' War began in 1618': the meaning of this proposition cannot be determined simply by reference to the start of the Thirty Years' War, but must be evaluated by subsequent events. Narrative sentences do not refer in the same way as propositions such as 'snow is white', but instead derive their meaning from the larger temporal structure of which they are a part (Danto 1965: 8).[2]

According to Danto, a historical explanation consists of a sequence of such narrative sentences $(p_1, p_2, \ldots, p_n)$ with a certain contrastive structure. In particular, Danto proposes that the form of an explanandum can be represented as: x is F at $t-1$ and x is G at $t-3$. The other narrative sentences involved in the historical explanation then serve to explain this change. In this way, the basic structure of historical explanation is:

(1) x is F at $t-1$

(2) H happens to x at $t-2$

(3) x is G at $t-3$

[2]It is this structure that substantive philosophies of history disobey: "There is a certain sense in which we can tell only *true* stories about the *past*. It is this sense which is somehow violated by substantive philosophies of history." (Danto 1965: 11)

Further, (1) and (2) do not necessarily entail (3): the explanatoriness of historical explanations is entirely independent of any predictive power it may or may not have. It also has nothing to do with Hempel's 'method of empathic understanding' – the method by which the historian supposedly "...imagines himself in the place of the persons involved" (Hempel 1965: 239). Instead, historical explanations carry the indelible mark of the particular perspective of the historian who, looking back over time, selects the events that are deemed significant to explain certain historical changes that are of interest to him or her, in order to create a story or 'narrative'.

This lack of predicative power casts doubt on the scientific character of historical explanations. This is a consequence Danto was happy to accept: he saw narrative explanations as closer in nature to art and fictional narratives than to scientific explanations, and wished distinguish sharply between narrative and scientific explanation. However, others were less happy to accept a dichotomy between narrative and scientific explanation, especially given the historical character of some scientific disciplines.

One of those who wished to avoid a dichotomy was Hempel, who tried to reduce narrative explanations to his deductive-nomological (D-N) model of explanation, formulated together with Oppenheim in 1948. In this formulation, the structure of an explanation is conceived as a *deductive argument*, where a given empirical phenomenon (the explanandum or E) is deductively inferred from the conjunction of general laws (L_i) together with certain conditions (C_i). These conditions identify certain relevant facts concerning the initial conditions that guide how the laws are to be applied. Schematically:

$$\begin{array}{l} C_1, C_2, \ldots, C_k \\ \underline{L_1, L_2, \ldots, L_r} \\ E \end{array}$$

Part of the intuitive force behind the deductive model of explanation is that it clarifies why explanation and prediction seem to go hand in hand. If A explains B, then knowing A should allow B to be predicted.

Many explanations in physics recognizably follow this pattern. For example, in Newtonian gravitational systems, only two laws are needed (the gravitational law and the second law), and the conditions C_i specify

the masses, and initial positions and momenta of the particles of the system. Given these two sets of premises any later state of the system can be deductively inferred. This explanatory structure is repeated in generalized classical mechanics and quantum mechanics (where the laws are, respectively, Lagrange's equations and the Schrödinger equation).

When Hempel turned towards the subject of explanations of history – both natural history and human history[3] – he attempted to make such explanations fit the mould of D-N explanation:

> Historical explanation, too, aims at showing that the event in question was not "a matter of chance," but was to be expected in view of certain antecedent or simultaneous conditions. The expectation referred to is not prophecy or divination, but rational scientific anticipation which rests on the assumption of general laws. (Hempel 1965: 235).

Curiously, Hempel recognizes but is puzzled by the fact that not only do there not seem to be any general laws in history, but historians seem to deny even their possibility:

> If this view is correct, it would seem strange that while most historians do suggest explanations of historical events, many of them deny the possibility of resorting to any general laws in history. (Hempel 1965: 235)

To resolve this apparent contradiction, Hempel introduces the notion of *explanation sketch* (Hempel 1965: 238). Historians do not offer full D-N explanations, but rather *sketches* of such explanations, where the general law and initial conditions are only vaguely or incompletely stated. For example, a historian attempting to explain why Dust Bowl farmers migrated to California would invoke the drought and the better living conditions in California. According to Hempel, this is a sketch in need of "filling out" (*Ibid.*: 238) in order to become a "full-fledged explanation": for example, it could be filled out with the universal hypothesis that "populations will tend to migrate to regions which offer them living condition"' (*Ibid.*: 236).[4]

[3]Human history is of course a natural phenomenon, but, because of free will, perhaps not a merely natural phenomenon analyzable by scientific investigation.

[4]Why then do historians not do the filling out for us? Hempel gives two reasons.

2 The Causal Approach: Between Nomothetic and Narrative Explanation

Hempel's notion of an explanation sketch is an awkward attempt to bring historical explanations into the D-N fold – awkward because the type of universal laws that Hempel would need for the D-N logic are simply unavailable in disciplines such as biology or social science (see chapter 4). However, philosophy of science moved away from the D-N model for other reasons, and has largely (though by no means universally) adopted a causal approach, which is, as we will see, a much more natural general framework in which to situate both narrative explanations in science, as well as interpretations of evolutionary history.

Ironically, the main counterexamples to the D-N model of explanation all involve some feature typical of narrative explanations. The three most well-known problems (which are stated here without much explanation) are (1) non-generality of the explanans in some explanations, (2) directionality in some explanations, and (3) criteria of 'relevance' in selecting the explanantia. An example of (1), offered by Scriven (1962), is the explanation 'the vase fell because it was knocked down by the curtains'. There is no general law concerning curtains and vases; moreover, such an explanation does not even implicitly involve universal laws, such as the laws governing electrostatic forces. Instead, according to Scriven, this type of explanation involves 'singular causation' rather than a general law doing the explanatory work.

An example of (2) lies in how we intuitively judge the length of a pole to explain the length of the shadow cast by the pole, but not vice versa – even though the length of the shadow cast deductively entails the length of the pole. Finally, concerning (3), if a male takes a birth control pill and fails to get pregnant, we judge his maleness to be relevant and not his taking a birth control pill for explaining why he fails to get pregnant – even though both entail the explanandum according to D-N criteria.

The first is that historians rely on common-sense hypotheses relating to individual and social psychology. The second is that it would be 'very difficult' to formulate the tacit hypotheses in a way that would be both precise and in agreement with relevant empirical evidence. However, Hempel never delves into why it should be difficult for historians to formulate such tacit universal hypotheses – and into whether this difficulty is a sign that such tacit hypotheses might not exist.

Given these problems, causality became an increasingly attractive notion with which to analyze the structure of scientific explanation (instead of universal laws). Causality had been largely repudiated by the logical empiricists (with some exceptions, such as Hans Reichenbach) as causality was thought to be a mere metaphysical construct, unnecessarily interpolating empirical data. For example, at one point Russell pronounced causality to be "a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm" (Russell 1912:1).

Nonetheless, causality deals fairly naturally with the problems mentioned above. Causation runs from cause to effect, so a causal explanation of the length of the shadow will cite the length of the pole, but not vice versa. Further, irrelevant explanantia are simply those that make no causal difference to the explanandum: there is no causal path that connects taking a birth control pill to a male getting pregnant. Finally, the causal account of explanation integrates some of the virtues of the D-N account: while a cause may not always be sufficient for the explanandum to occur, it does allow for a certain degree of predictability (often in terms of 'statistical relevance': see Salmon 1971).

There have been numerous accounts of causal explanation (Salmon 1971, 1984; Dowe 2000; Woodward 2003; Strevens 2008), and a proliferation of notions of causality (the main ones being causal production and causal difference-making[5]). Old Humean concerns have never entirely disappeared: for example, it is an ongoing concern how to distinguish causation from correlation, or whether causal relations are simply regular patterns of succession[6].

[5]Causal production refers to spatiotemporally contiguous processes that connect events (Salmon 1984, Dowe 1990). A problem with this is that for many explananda, there are many processes causally linked to the explanandum without being explanatorily relevant. If Joe kicks a football at a window and the window shatters, mentioning that Joe was at the time performing a difficult curve ball seems superfluous to explaining why the window shattered – his having kicked the ball is sufficient. Not all causal detail is relevant, and this is why a number of the most prominent accounts of scientific explanation are spelled out in terms of causal difference-making (Woodward 2003, Strevens 2008).

[6]An influential approach (among philosophers as well as computer scientists) has attempted to resolve this problem by *defining* causality in terms of certain types of correlations that can be cleanly *manipulated* by surgical interventions (Spirtes et al. 2000; Pearl 2000; Woodward 2003).

We need not go into further detail; some contemporary issues in the literature on causality will be discussed in the context of the causality of natural selection in chapter 5. What is more important to note here is that the main motivations for a causal approach – the specificity, directionality and selectiveness of explanations – are precisely the features that Danto referred to in order to distinguish narrative explanation from law-based predictive explanations. First, historians are concerned with idiosyncratic detail, not with generalities; second, due to the chronology of events, narratives have a clear directionality; and finally, narratives do not represent the totality of past events (the Ideal Chronicle), but make a selection.

Explanations that use the concept of path-dependence can be seen as further elaborations of the causal approach to explanation. Instead of the explanans being a single cause (or a pair of contrastive causes), path-dependent explanations will cite a causal pathway (a series of antecedent causes) and will contrast that pathway to pathways that could have been taken. In what follows I will give a broad overview of phenomena that are usually explained as path-dependent. This will allow the challenges of description-dependence to be formulated as challenges for an account of path-dependent explanation.

3 The Phenomenon of Path-Dependence

Path-dependent phenomena occur across all domains of inquiry, from the study of topological defects in solids to the adoption of technologies. A paradigm case of path-dependence is often taken to be the adoption of the QWERTY keyboard (David 1985). Originally, the QWERTY layout was designed so as to to prevent typewriters from jamming; however, it subsequently became entrenched, even though the typewriter itself became obsolete, and even though there are more efficient ways of organizing an English-language keyboard for a computer. Thus, to explain why the present state of keyboards is as it is, one needs to integrate information about past states.

In the broadest sense, path-dependence merely implies that the path followed by a system is explanatorily relevant for its final outcome. In this sense, the term is simply another way of saying 'past states matter'. However, once this 'explanatory relevance' is given a more precise char-

acterization, narrower and more technical accounts of path-dependence emerge. Such accounts were originally proposed in economics and the social sciences (Arthur 1994; Pierson 2004), but more recently the issue has been receiving increasing attention from philosophers of science (Szathmáry 2006; Ereshefsky 2012; Desjardins 2011a,b, 2015).

In the philosophy of science literature, the notion of path-dependence is closely related to two other ways of understanding historicity: 'sensitivity to initial conditions' (Ben-Menahem 1997; Powell 2012) and 'contingency' (Beatty 2006; Beatty and Desjardins 2009). Contingency is particularly important for the purposes of this dissertation, and while an outcome can be contingent without being path-dependent (see chapter 2), such outcomes are of negligible importance for evolutionary history. The precise relation between path-dependence and contingency, and between path-dependence and sensitivity to initial conditions will become clearer over the course of this and next chapter.

A description of some salient properties of path-dependent explanations of processes[7] will help develop some intuitions concerning the phenomenon. I will loosely group these properties according to whether they are future-oriented or past-oriented.

Among the past-oriented aspects, a key distinction is that between information-preserving and information-destroying processes (Sober 1983, 1988; Desjardins 2011b). The latter is exemplified by what happens when a marble is released from the rim of a bowl: the marble will roll down and rock back and forth until it comes to a stop in the middle of the bowl. Given information about the final state alone, it is impossible to reconstruct its initial state, since no matter where precisely on the rim the marble is released, it will invariably come to rest at the middle point. This is an example of path-independence, as the precise path followed by the

[7] I primarily take path-dependence to be a property of an *explanation* or *representation* of a process, not a property of the process itself. While some processes lend themselves more naturally to a path-dependent representation than others, as we will see processes can be represented as both path-dependent as well as path-independent according to the grain of analysis adopted (this is in fact the whole problem of description-dependence). Hence attributing path-dependence to an explanation is the only way to avoid ambiguity. For the sake of brevity, I will often refer to representations of processes simply as 'processes', and explanations that represent processes as path-dependent as 'path-dependent explanations'.

marble makes no difference to the final state. In other words, the past of the marble is 'erased' and does not matter for the explanandum.

One of the most basic path-dependent processes is movement with friction. If one slides a block of wood from point A to point B, then it matters whether the shortest path between the points is chosen, or some more indirect route. In the latter case, more heat will be generated due to friction between the block of wood and the surface. Thus, some information about the past (i.e. the length of the path followed, or the speed with which the block was pushed) is preserved in the final state.

In general, most real processes have both information-preserving and information-destroying aspects, and in this way can only be said to be partially path-dependent. For example, the morphology of the whale shows remarkable similarities with the morphology of fish, yet there are significant differences as well. Some information about the past is destroyed due to the convergent evolution towards the streamlined morphology. However, a whale has lungs instead of gills, and its fins are exapted from fingers, and thus some information about its land-based past is preserved.

A second group of properties of path-dependence concerns how the past makes a counterfactual difference for the present: if the past were different, the outcome would also be different. For example, if humanity had skipped the technology of typewriters, and gone straight to computers, there would likely be no QWERTY keyboard. The phenomenon of sensitivity to initial conditions — how a small change in initial conditions can lead to a large change in outcomes — concerns this aspect of path-dependence (Ben-Menahem 1997; Powell 2012). An example is the nonlinearity of the weather, so that, so to speak, a butterfly can flap its wings in Paris and cause a storm in New York. The outcome could not have occurred if the past were different.

An underlying notion here is the contingency of the explanandum, where 'contingency' refers not to the modal structure of the explanandum (i.e. whether or not it is true in all possible worlds), but to the structure of its causal history. The outcome is contingent if, given what we know about its causal antecedents, it could not have occurred.[8]

[8]This is partially why historical explanations do not fit the mould of deductive (or even inductive) explanations. The explanandum cannot be deduced from a general principle, or inductively inferred with high probability, but maintains some degree of

It is helpful to distinguish between two types of contingency: 'causal-dependence contingency' and 'unpredictability contingency' (Beatty 2006). Causal-dependence contingency refers to the counterfactual dependence of the outcome on some prior state. Thus A is 'contingent upon' B if, and only if, were B not present, A would not obtain. Causal-dependence contingency is thus a very broad notion, and also covers deterministic processes where there is dependence on initial conditions, such as the Newtonian dynamics of individual particles.

Unpredictability contingency refers more specifically to indeterminism in a process, or at least, a modelled indeterminism in the explanatory structure[9]. It is insufficient to know the prior states in order to 'predict' the outcome state. Beatty describes this as 'contingency per se' (Beatty 2006: 38-40), thus indicating that contingency can also be used as a one-place predicate attaching to an explanandum.

These two notions of contingency capture two different senses in which the outcome could not have occurred, given the initial state. Unpredictability contingency is not necessary for path-dependence, as some perfectly predictable processes (e.g. Newtonian dynamics) depend on initial conditions, and thus the outcome is dependent on which causal path had been taken[10]. This has led some to refer to dependence on initial conditions as 'weak' path- dependence, and the cases where the process depends on multiple past states as 'strong' path-dependence (Ereshefsky 2012).

These two orientations, future-directed and past-directed, are in no way mutually exclusive, and most real examples involve both perspectives. Consider the phenomenon of positive feedback, where the system is initially balanced between two basins of separate attractors, and where any initial fluctuation will snowball and result in a large, self-reinforcing change. A classic example of this is the emergence of the VCR videocassette technology (Arthur 1994). Initially, the videocassette market was precariously poised between two competing technologies, VCR and Beta; however, a slightly greater adoption of the VCR technology by consumers led to it becoming more widely available in video outlets, in turn precipitating fur-

'contingency'.

[9]Many processes in statistical physics and the special sciences are modelled as probabilistic, even though the underlying causal processes may be deterministic.

[10]We will see later that unpredictability contingency is not sufficient either: some probabilistic processes are ahistorical.

ther adoption by consumers. Thus VCR came to dominate the market. In this example, there is unpredictability contingency (the initial greater adoption was purely contingent) and sensitivity to initial conditions, but also some information-preservation, as given the outcome of VCR dominance, we can extract some information about some of the past states.

Another interesting combination of both orientations occurs in instances where the initial state neither snowballs nor is erased, but where it simply constrains future evolution. For example, developmental mechanisms, such as the processes determined by the Hox genes, constrain possible body-plans and thus the adaptations that are possible (Young et al. 2005). There is a counterfactual dependence in the sense that past states (like a certain configuration of the Hox genes) preclude some possible future states. When the past constrains the outcome to the extent that only a single outcome becomes possible, the phenomenon is known as 'entrenchment' or the 'lock-in effect'. There is also some information-preservation here, as it is possible to reconstruct the past to a certain extent.

4 Description-dependence: Three Challenges

With these phenomena in mind, three challenges face any account that attempts to uncover the more formal structure of path-dependence. The first is to account for how path-dependence is a matter of degree. While some measures have been intuitively suggested in the literature (e.g. Desjardins 2011a), a more rigorously developed account is lacking. This is partly due to the fact that the literature is relatively new. However, perhaps it is partly due to some confusion about two ways in which 'degree' can be understood.

The first way is when the past matters at multiple moments instead of a single instant. Thus, insofar as the evolution of the whale is represented as depending on at least two moments (the transition from fish to land-based animal, and from land-based to aquatic reptile) instead of just one in the case of VCR history (the instant when VCR happened to become more frequent than Beta), the evolution of the whale can be considered more path-dependent than that of the VCR. The outcome state gives more 'information' about the past in the first case than in the second.

The second way the past matters 'more' is when a difference in the past

leads to a 'greater' difference in outcome. For example, the past matters 'more' when a butterfly flapping its wings leads to a hurricane than when the flapping would merely lead to a small displacement of air. Note that these two types of degree are not necessarily equivalent: for example, if the hurricane happens to be some attractor state, in such a way that many other kinds of small disturbance would likewise have led to the hurricane, the fact of a butterfly flapping its wings is not very relevant to explain why the hurricane occurred. The background conditions (pressure, temperature differentials) would be more informative; what actually triggered the hurricane would be relatively unimportant. In this way, while the first measure of path-dependence concerns how informative the past is for explaining the present, the second measure compares the 'distance' between the outcome states of given initial states.

I will be leaving this second sense of degree aside, mainly because the information-focused sense of degree is more fundamental and leads to interesting connections with information theory. However, another reason is that formalizing this second sense of degree would not be worthwhile for the purposes of this dissertation. Allow me to briefly sketch why. The distance-focused degree can either refer to nonlinearity or discontinuity.[11] If one takes it to be discontinuity, it is a discrete property of a process and hence not a good candidate for a gradualist degree of path-dependence. If one interprets it as nonlinearity, then one would need to detail what it means for one outcome to be 'very different' from another. Which metric is one to use, for example, on the space of possible videocassette technologies? It seems impossible to introduce any metric without relativizing it to explanatory interests. Thus, if taken as nonlinearity, this distance-based degree of path-dependence seems to mainly depend on what explanatory interests are at stake rather than on the nature of path-dependence.

Besides accounting for how path-dependence comes in degrees, a second challenge is that the evolution of a system may be path-dependent at only

[11] In brief, a function is linear when $f(x + y) = f(x) + f(y)$; thus when a function is nonlinear a slight change in input will lead to an effect that is not linearly proportionate, and could potentially be very large. When a function is discontinuous, some modifications of the input, no matter how slight, will lead to relatively large effects. If a process is nonlinear but continuous, small changes will still lead to small effects; however, in a discontinuous process, some changes, no matter how small, will lead to large effects, even if the process is otherwise linear.

certain times, or only with regard to certain outcome states. Thus path-dependence seems to have different scopes, some more local, others more global.

A third, final challenge concerns the way in which path-dependence seems to depend on the grain of analysis adopted to describe the process. For example, the evolution of the whale is path-dependent when one distinguishes between the two states 'fish' and 'marine mammal'; however, path-dependence disappears when the outcome state is more coarsely described (e.g. 'aquatic animal'). What counts as an adaptation or a constraint is to some extent dependent on the grain of analysis (see Wilkins and Godfrey-Smith 2009). In general, introducing a more fine-grained description of the explanandum seems to make it more path-dependent, and an account of path-dependence should be able to integrate this fact.

Illustration: Homoplasy and Homology

Particularly relevant for our purposes is how to disentangle the relative roles that natural selection, development processes and genes play in causing two lineages that share a similar trait. If the similarity in the trait is due to common descent, the similarity is said to be *homologous*[12]; if it is derived 'independently' (not due to common descent, but for instance, due to natural selection), it is said to be *homoplastic*.

However, traits can be described at different levels of analysis: at the levels of function, structure, underlying developmental processes and underlying genes or gene networks that ultimately code for the phenotypic trait. Two traits in separate evolutionary lineages may be similar at the level of function while being dissimilar at the level of structure, or similar at the level of structure while possessing dissimilar developmental processes.

In this way, the question, 'Is the similarity due to common descent or not?' is ambiguous as such, and admits of different answers. For example, there is often some common descent at play even in cases that are cate-

[12]Some (e.g., Ramsey and Peterson 2012) have defined homology in terms of sameness instead of similarity, since similarity leads to a number of definitional ambiguities. However, also in Ramsey and Peterson's account of homology, description-dependence remains an issue, since they distinguish between three levels of analysis: the genes coding for the developmental programme, the developmental programme, and the product of the programme, i.e., the morphological trait itself (Ramsey and Peterson 2012: 262).

gorized as convergent evolution. A well-known example is how the Pax-6 gene is present in the developmental programmes controlling for both vertebrate and cephalopod eyes; yet because the developmental programmes are so different, the similarity between vertebrate and cephalopod eyes is usually categorized as convergence. At the level of genes there is some common descent; at the level of phenotypic structure, similarity is not entirely explained by the common Pax-6 gene (even though without Pax-6 there would likely be no similarity).

In general, the phylogeny (causal network) that traces the evolution of a trait across different lineages is complex and highly description-dependent. This complexity allows for different permutations of convergence and divergence at the different levels, including for example: 'convergent evolution' (homoplasy at the level of structure without homologous developmental programmes), 'parallel evolution' (homoplasy with homologous developmental programmes), or 'deep homology' (no homoplasy at either the level of developmental programmes or structures, and yet homology at the level of genes: see Ramsey and Peterson 2012). Another consequence of this complexity is that the labels do not necessarily cut the causal network 'at the joints'. For example, there may be similarity at the level of a structure, but the developmental programmes may only be partially homologous.

Conway Morris uses the ubiquity of homoplasies in order to argue that convergent evolution is ubiquitous – and hence that many evolutionary outcomes are inevitable. However, the description-dependence of traits means that in practice it is difficult to establish that a homoplasy is the result of convergent evolution and not, for example, parallel evolution. Powell (2012) uses this fact to argue against the convergence-centric interpretation of evolutionary history, and argues that many instances of homoplasy, being the result of homologous developmental programmes (such as the programmes controlling for body plan), actually support the Gouldian interpretation of life.

In chapter 3 (and in part II in more detail) we will see how Powell's analysis of convergent evolution, while not wrong, is incomplete and overlooks the fact that even convergent evolution does not support the case that evolutionary outcomes are non-contingent.

Chapter 2

The Symmetry Framework

In this chapter we argue in more detail why various types of contingency should be analyzed in terms of path-dependence, and in particular in terms of symmetries of causal networks. The specific approach in this chapter will be to follow recent developments in the philosophy of science and to analyze narrative explanations in terms of path-dependence – this idea is not new, and has already a large following in historical sciences such as political science and sociology, but has only recently been introduced into the philosophy of science. I will present a novel understanding of narrative explanation in terms of symmetry-breaking.

While this endeavour can be pursued as a stand-alone project within the literature on path-dependence (see Desmond 2016), the purpose of this dissertation will be to find the tools with which a more nuanced hypothesis concerning evolutionary contingency can be formulated – more nuanced than the hypotheses associated with the Chardinian and Gouldian narratives.

1 The Attractor Landscape and Branching Tree Models

Among the more technical accounts of path-dependence, two classes of model can be discerned. The first characterizes path-dependence as occurring when a system could possibly evolve towards one of multiple local stable equilibria or *attractor states* (*e.g.* Bassanini and Dosi 1999;

Pierson 2004; Szathmáry 2006). In this way certain key aspects of path-dependence, such as nonlinearity and sensitivity to initial conditions, can be modeled as an evolution on what I call an *attractor landscape* with multiple attractor states. Such a model, like the adaptive landscape metaphor in evolutionary biology, has serious limitations, the main one being that in systems with high dimensionality the topology of associated landscapes tend to be ridged and holey. As we will see, this means that the dynamics of such systems cannot be modeled as simply maximizing some scalar variable, and this precludes (or at least seriously limits) a general formulation of path-dependence in terms of landscape topology (see Gavrilets 2004; Kaplan 2008).

The second broad class of model has represented path-dependent processes as a *branching causal tree* (Kaplan 2008). However, branching trees also have limitations when the causal structure becomes too complex, in particular when there are multiple possible initial states, or when there is a significant number of non-tree events, or 'reticulations', where branches converge (Moret et al. 2004). For example, an area where the tree metaphor has received significant criticism is phylogenetics, where phenomena such as hybrid speciation or lateral gene transfer cannot be captured in tree models.

To address these limitations, the first purpose of this chapter is to introduce the notion of *causal networks* in some formal detail, and show how they are generalizations of both landscapes and branching trees. Network models are already well established in the causal modeling literature (following Pearl 2000) and in phylogenetics (*e.g.* Moret et al. 2004; Velasco and Sober 2010), but in the literature on path-dependence they have been underutilized. Causal networks allow complex causal relations to be represented when both tree or landscape metaphors fail.

An attractor is an equilibrium set of states towards which the system evolves when it is in a given neighbourhood (the 'basin'), and once in the attractor state, the system will tend to return there if perturbed. Its usefulness as a concept primarily lies in allowing for some long-term predictability, even in dynamics that are nonlinear and chaotic. An attractor is *global* when its basin covers all of state space, or *local*, when the basin is a subset of state space. In what I call an 'attractor landscape', each state is assigned a scalar variable (on a two-dimensional landscape, this

is the height), with the attractors being local maxima (or minima), and the system tending towards maximizing (or minimizing) the scalar variable. Examples of such landscapes are potential energy landscapes, where valleys in the landscape correspond to minimal-energy states, or adaptive landscapes, where the peaks represent states with maximal fitness.

Landscapes can be used to systematize some aspects of path-dependence, for example, the distinction between information-preserving and information-destroying processes. Reconstructability becomes impossible when the explanandum (the outcome state) is a global attractor state, because any possible initial state tends towards the attractor state. When there are multiple attractors present, the process is partially information-preserving, as one can extract some information about the past (namely, in which basin the system was initially located) from the outcome.

With this in mind, one could formulate path-dependence in terms of the following negative condition:

Definition (Path-dependence - attractor formulation)**.** *An explanation of an outcome is path-dependent if and only if that outcome is not explained as a* ***global attractor****.*

Note that 'global' is always defined relative to the state space under consideration. The middle of the bowl is a global attractor when the state space is confined to the positions of the marble on the hollow surface of the bowl; it (obviously) is no longer an attractor when the marble is placed next to the bowl. Thus, when an attractor is deemed global within the scope of the explanation, then what precise initial state obtains does not make any difference for the outcome, as the system will be in the attractor state. Conversely, when there is no global attractor, then there are at least two initial states that lead to different outcomes.

The accounts of Bassanini and Dosi (1999) and Szathmáry (2006) implicitly draw on this criterion. Szathmáry distinguishes between 'strong' and 'weak' path-dependence (not to be confused with Ereshefsky's distinction). Strong path dependence occurs when the process is irreversible and when there are multiple stable attractors. This is straightforwardly covered by the attractor formulation.

However, what Szathmáry calls weak path-dependence could seem problematic for this definition. An outcome may not be a global attractor, and

yet have occurred path-independently in the weak sense, as long as the causal-dependence on initial conditions is 'effectively' eliminated as time goes to infinity.[1] This type of weak path-dependence will tend to occur in high-dimensional state spaces, when the number of possible states is 'much' greater than the number of states actualized over the course of a system's history, so that the asymptotic convergence of possible trajectories will tend not to occur (Szathmáry 2006).

Nonetheless, weak path-dependence is also covered by the attractor formulation, because the asymptotic convergence that Bassanini and Dosi describe concerns the convergence of the *average* position of a trajectory. Even though the actual *instantaneous* positions of two possible trajectories will in general be very different at any given time, when a system is weakly path-dependent, the long-run average position converges to a global equilibrium state.[2] Thus, whether or not the past matters in weak path-dependence depends on what precisely the explanandum is: the average position over a long period of time (past does not matter), or the actual position at a specific time (past does matter). By contrast, strong path-dependence implies that both the instantaneous and the long-run average position converge to a single attractor state.

The landscape framework has serious limitations. An area where it has already received significant criticism is in its application to biological evolution in the short- to middle-term (*i.e.* adaptive landscapes).[3] One important criticism concerns how landscapes change when the dimensionality of state space is increased. Landscapes imply that a system

[1] Effective elimination is what Bassanini and Dosi (1999: 15) call asymptotic path-independence, which occurs when two possible trajectories come arbitrarily close within a finite time-span, and for an infinite number of times thereafter. (If the dynamics is Markovian, then this condition reduces to the following: two possible trajectories intersect in finite time, because once there is a single intersection, it is expected that the paths will overlap for all subsequent times.) If this condition is met, then the difference an initial condition makes on a subsequent history is eliminated in finite time. In this way, weak path-independence is a form of ergodicity.

[2] Compare with Doeblin's theorem in the theory of Markov processes (*e.g.* Stroock 2005).

[3] Note that while some have argued in their defence that they are best used as an explanatory template, as a heuristic for hypothesis generation (Ruse 1990; Skipper 2004), others have questioned their adequacy even as metaphor (Pigliucci and Kaplan 2006; Kaplan 2008; Plutynski 2008).

can evolve smoothly to a neighbouring state; Gavrilets (2004) has shown how the topology of high-dimensional adaptive landscapes tends to consist of'ridges', 'rugged peaks' and 'holes' rather than of smooth hills. The likelihood increases of a (nearly) neutral network forming, and of the number of local peaks increasing (see Gavrilets 2004: 45-80). The absolute scalar difference (in this case, fitness) between any two states becomes increasingly meaningless for predicting whether one state will evolve into the other or not.

What this suggests is that according as one needs more variables to characterize a particular outcome, the less likely one will be able to analyze the occurrence of that outcome as some kind of optimum of a single scalar quantity (*e.g.* fitness). While the attractor formulation of path-dependence may remain true, it becomes increasingly empty, as simple global attractors tend to not occur at all in complex systems. As landscapes become increasingly ridged and holey, the basins of local peaks shrink, and the system likely does not exhibit any global optimizing behaviour.

In this way, attractor landscapes may have limited applicability. Furthermore, they cannot represent many interesting path-dependent processes; they are best suited to represent convergent processes, or processes where there is a choice between multiple local attractors.[4] Causal trees are better suited for evolutions in high-dimensional spaces, where the probability that causal paths intersect is very low, and thus where every state actualized is unique.

1.1 Causal Trees

In the following I will briefly outline a formal characterization of trees, and then (drawing on the work of Desjardins) consider how path-dependence can be formulated within this framework. I will try to show that this framework is in a sense the opposite of attractor landscapes: best suited for high-dimensional state spaces, but weak at representing convergent causal structures.

A **tree** is a causal graph rooted in a single point, from which branches

[4]Compare this with the analysis of conservative vector fields: if a dynamics can be represented as the gradient of a scalar, then it is path-independent.

split off but never join as one moves from past to present. The states of a tree form a partially ordered set of states, where every state has only a single immediate predecessor, but can have any number of successors. Thus not every pair of states can be connected by a forward-directed causal chain, even though every state in a tree is indirectly causally linked through a common ancestor.

A causal tree maps out the possible paths an individual entity can follow. If the system consists only of a single entity, only a single path will actually be followed; if the system is an ensemble of individual entities, there will be a distribution over the possible paths according to the probabilities of the paths. The branching events or *nodes*, which connect a single state to two or more possible descendant states, can be thought of as abstractions of *contingent events* with causal impact on the path of the system. For example, in macroevolutionary phylogenetic trees the nodes abstractly represent speciation events, where a given biological population diverges to two or more distinct species.

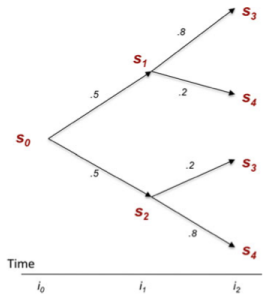

Figure 2.1: Path dependence in a causal tree (Reproduced with permission from Desjardins 2011a)

With this in place, one can formulate path-dependence in the following way (adopted from Desjardins 2015)[5]:

Definition (Path-dependence - causal tree formulation)**.** *An explanation of an outcome is path-dependent if 1/ a given initial state branches off into at least two paths, 2/ these paths lead to at least two possible outcomes (with non-zero probability), and 3/ following different paths affects the probability of a given outcome state.*

[5]For a more mathematical characterization, see Desjardins (2011a).

This formulation of path-dependence captures some crucial properties, such as unpredictability contingency and causal-dependence contingency. It can also be used to capture the information-preserving aspect of path-dependence, and the way in which it can come in degrees (Desjardins 2011a). However, I would like to point to three limitations. The first and foremost is that, in contrast to attractor landscapes, causal trees cannot capture causal relationships where branches join. This is a problem for even the formulation of path-dependence, as path-dependence presupposes that there are alternative paths leading to the same outcome, and thus some convergence. This can be seen more clearly by redrawing Figure 2.1 so that the same states are represented by the same points; then the causal model becomes Figure 2.2, which, strictly speaking, is no longer a causal tree.

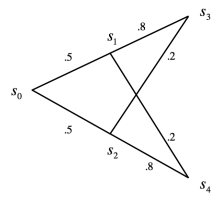

Figure 2.2: Path dependence in a causal network

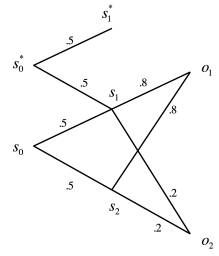

Figure 2.3: Path-dependent or not?

Putting this problem aside (for example, by expanding the notion of tree to allow for some reticulations), it remains unclear how to analyze cases with multiple possible initial states. For example, in Figure 2.3, none of the paths leading to o_1 affect the probability of o_1 occurring, and thus the occurrence of o_1 does not seem to be path-dependent in the sense that its occurrence is not affected by the choice of path. Yet, there is a clear dependence on initial conditions, for if one knows that the initial state is s_0^*, the probability of o_1 occurring is .4, as opposed to .8, if s_0 were to be the initial state. The example in Figure 2.3 thus seems to involve some combination of path-dependence and path-independence that is not captured by the causal tree formulation.

A second, related limitation is that the causal tree formulation concerns only whether the occurrence of an outcome is path-dependent, but it is unclear how it can be applied to a set or distribution of states, or how path-dependence is something that can change over time. In other words, the tree formulation does not seem to allow for different *scopes* of path-dependence. For example, in a more complex model such as Figure 2.4,

there seem to be pockets of path-independence, even though the process may be globally path-dependent. The origin of this limitation lies less in the specific formulation of path-dependence, but rather in the causal tree framework itself; this is one important reason for representing causal relationships by causal networks (directed acyclical graphs).

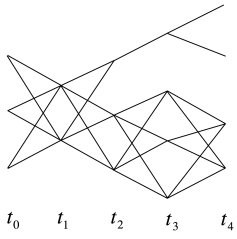

Figure 2.4: Path-dependent or not?

Finally, it remains unclear how precisely the degree of path-dependence should be defined. Desjardins (2011a) suggests two types of metric that roughly correspond to those mentioned in section II. The first is the degree of divergence or convergence, where maximal divergence is maximal path-dependence and maximal convergence is maximal path-independence. The second is the degree of 'similarity' between outcomes: a causal tree is more path-dependent when the different outcomes are more dissimilar. However, it would be desirable to introduce a more precise, quantitative measure.[6] It is not clear, within a causal tree, what 'similarity' between outcomes could mean without introducing some independent scalar metric.

1.2 Causal Networks

As done with causal trees, I will now construct causal networks with some more formal detail. Besides allowing for increased generality when describing path-dependence, there are two further advantages in doing this. The first is that it will become clear how a model can be *coarse-grained* to obtain either a causal tree or an attractor landscape, thus showing how the two frameworks are limiting cases of causal networks. The second is that it places path-dependence within the context of graph theory, to which the tools of information theory can be readily applied, and this will allow for a quantitative measure of path-dependence to be proposed.

A **causal network** is a directed, acyclical graph represented by the ordered pair (V, E), where V is the set of nodes and E is the set of edges connecting the nodes. In this dissertation, causal networks are taken to

[6] Also, it can be shown that maximal divergence is, perhaps surprisingly, a case of maximal path-independence (see Figure 2.8).

be formalizations of causal explanations, and hence certain nodes are of particular interest, namely the outcome states and the initial states. For this reason it will be useful to think of the ordered pair (V, E) as a 3-tuple (O, I, R) where O is the set of outcome states, I the set of initial states, and $R : O \rightarrow I$ a web of causal relations between initial and outcome states. The causal relations themselves may be productive or difference-making — the precise nature of causality will not be of concern here. In general, causal networks will contain intermediate states, between the sets of initial and outcome states. Letting these intermediate sets of states be represented by $O_i = I_{i+1}$, with $I = I_0$, $O = O_n$, the relation R can be decomposed in $n + 1$ instants: $R = R_0 \circ R_1 \circ \cdots \circ R_n$, where each $R_i : I_i \rightarrow O_i$ is a simple mapping relation.

Three basic causal patterns will be of interest. In a **parallel** structure, the outcome would not have obtained if a particular initial state had not been present. Thus, there is at most one initial state associated with a given outcome, and in this way the parallel structure corresponds to causal-dependence contingency. By contrast, in a **divergent** structure, multiple outcomes are associated with a single initial state. This means that, given the initial state, the descendant state cannot be predicted: this is unpredictability contingency. When the structure is neither parallel nor divergent, it is **convergent**, and this occurs when multiple initial states converge on a single outcome state. A path-dependent explanation, as actually used in scientific practice, is almost invariably a complex combination of these basic structures.

The **probability of an outcome** in a particular explanatory framework can be calculated by means of the probability distribution over initial states, and the probabilities of the different paths between an initial state and the outcome. By the law of total probability we get $P(o) = \sum_i P(i)P(o|i)$. Each conditional probability $P(o|i)$ can be written as $P(o|i) = \sum_p P(p_{io}) = \sum_p \prod_{i \leq j \leq o} \pi(j \rightarrow j + 1)$, where the p_{io} are the different paths connecting initial state i to outcome o, and where $\pi(j \rightarrow j+1)$ represents the transition probability connecting two past states. Thus the probability of an outcome is ultimately reducible to the initial probability distribution and the structure of the causal pathways leading to the outcome.

In general, the causal structure changes by fine-graining or coarse-

graining the degree of analysis. Fine-graining can be thought of as introducing a new variable to characterize the initial or outcome states, and in this way states that were previously identical become differentiated. More explicitly, a state may be characterized by n variables, $s = (x_1, x_2, ..., x_n)$, and one example of fine-graining is to introduce m new degrees of freedom, defining a new state $s' = (x_1, x_2, ..., x_n, ..., x_{n+m})$. So a single state s in n-dimensional space corresponds to an m-dimensional set in the fine-grained (n+m)-dimensional space. In this way fine-graining can be thought of as a one to many mapping, where a single state is replaced by a set of states.

The inverse operation is coarse-graining, and this is done by means of an *equivalence relation* $\sim$, which allows one to express that multiple states are 'similar' in some way. The equivalence relation defines an equivalence class on the states, $O/\sim$, where all the states that are 'similar' are represented by a single state. One way this can be done is by abstraction, where certain degrees of freedom are dropped, so that only the other features of a state are considered.

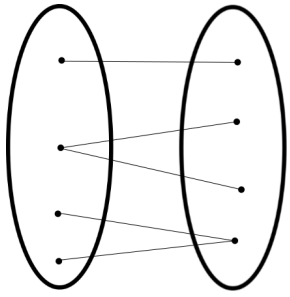

Figure 2.5: Parallel, divergent and convergent structures.

This offers a first step in making sense of how path-dependence is sensitive to the grain of analysis adopted in a causal model. Going back to the example of the evolution of the whale, what is striking here is that there is both convergence towards a fish-like morphology, and a divergence in other respects (such as bone-structure or respiratory system). One way to analyze this is the following: the evolutionary process can be represented as convergent evolution when the aquatic mammal state (AM) and fish state (F) are characterized by a single variable — their overall morphology; however, when the two states are characterized by additional variables (bone-structure, respiratory system, *etc.*), the evolutionary process is divergent. In the first case, the paths $F - M - AM$ and $F - F - F$ converge; in the second, fine-graining introduces path-dependence in the representation of the evolutionary process (figure 2.6).

One can summarize the effects of the grain of analysis on causal structure by means of the following (a proof is provided in the appendix):

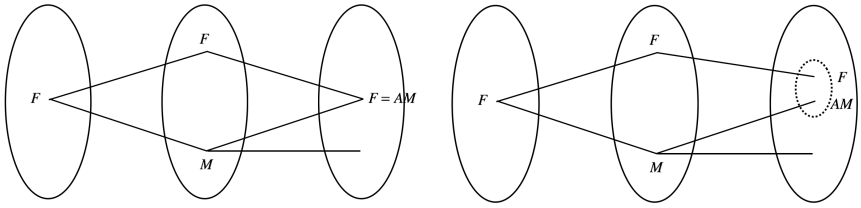

Figure 2.6: Two representations of the evolution of the whale. The right side representation models the evolutionary process in detail, and is path-dependent. The left side one coarse-grains over the aquatic mammal state (AM) and fish state (F), and represents the evolution of AM as (relatively) path-independent.

Theorem 1. *A coarse-graining of the explanandum makes an explanation increasingly convergent and a coarse-graining of the explanans makes an explanation increasingly divergent.*

This theorem gives some deeper insight into why attractor landscapes and causal trees are limited. In any attractor landscape, there is a countable number of privileged outcomes (attractor states), and each of these outcomes will have an associated subset of the initial states (the basin). When state space is described at a finer level of detail — *e.g.* when more variables are needed to adequately describe each state — the convergence of each basin on its respective attractor state will tend to decrease. A given attractor state will be disambiguated between two different states, each with its own basin. Ultimately, when the outcome states are described with sufficient detail, there will be no convergent structures any longer, only parallel structures, and the landscape metaphor disintegrates.

By contrast, the causal tree framework tends to be adequate as long as the number of possible states is much greater relative to the number of realized initial states, so that the probability of reticulations occurring is small. For example, this occurs when the dimensionality of the state space is relatively large. Taking the number of variables necessary to describe an entity to be a proxy for the complexity of that entity, this can also be formulated in terms of complexity. The dynamics of an individual, complex entity is likely to be path-dependent. By coarse-graining the state space (representing the complex entity abstractly) while keeping the

number of initial states constant, the convergence of the network increases monotonically, and the 'tree-ness' of the network decreases. In this way a causal tree can be seen as the limiting case of a causal network when the state space is much larger than the set of initial states.

2 The Symmetry Formulation

In this section I will propose how the concept of *symmetry breaking* can be used to characterize path-dependence and historicity in causal networks. The motivation for this proposal comes from the two main ways symmetry is used in physics (see also Brading and Castellani (2007)). The first, and most intuitive application of symmetries is to *properties* of a system, usually spatial configurations. A spatial configuration is symmetrical when it remains the same under some distance-preserving permutation of the elements (reflections, inversions and rotations). For example, a snowflake has some rotational symmetries (its appearance is unchanged when you rotate it by a multiple of 30°), reflection symmetries and a point symmetry. Similarly, a liquid has a maximal spatial symmetry: no matter how one would rotate, invert or reflect it, it would look the same. Such symmetry is *broken* during the transition to a solid: a particular molecular structure arises which will typically only have a limited number of symmetries.

Symmetries are also applied to the *dynamics* of a system, *i.e.* the way in which two subsequent states of a system relate to each other. Thus, instead of transforming the physical elements of the system, the variables in the laws of motion are transformed, and a symmetry is said to be present when the laws of motion remain invariant. In other words, the transformation is a symmetry of the dynamics if the transformed variables are related to each other in the same way as the untransformed variables are. One well known example is the time symmetry of Newtonian dynamics: because the second law gives a relation between the force and the second time-derivative of position (*i.e.* the acceleration), it is invariant under the transformation $t \rightarrow -t$. Thus, if one were to see an animation of a group of interacting particles, one could not tell by Newtonian dynamics alone whether the animation was being played forwards or backwards. In thermodynamic phenomena this time symmetry is broken: heat flows from warm to cold (the entropy increases), but never from cold to warm. A

rewound heat flow does not obey the second law of thermodynamics.

Here I will extend the notion of symmetry to the space of possible causal paths between past states and a particular outcome. A network will be symmetrical when the different past states can be permuted without affecting the causal structure of the network. Just as the snowflake remains unaffected by rotations, path-independent causal networks remain unaffected by permutations of past states (both initial and intermediary states). In itself, this basic idea is not much more than a reformulation of path-independence in the broad sense; however, it offers the resources to deal with some of the shortcomings of the tree and landscape frameworks.

2.1 Symmetry

More formally, let P_s be the probability distribution over the outcomes given that the system is in state s. One way to think of P_s is as the probabilities of the different possible outcomes as 'viewed from' s. The probability of any particular outcome o as viewed from s can be written as the sum of the probabilities of the different possible paths between s and o:

$$P_s(o) = P(o|s) = \sum_p P(p_{os})$$

where the variable p_{os} represents the possible paths between o and s. When there is only a single initial state s_0, one can assign an unconditional probability to an outcome $P(o) = P_{s_0}(o)$. This is the case in causal trees; however, in a general causal network, there is no unique way of specifying the unconditional probability of an outcome.

Note that these probabilities need not imply any fundamental indeterminism. For example, in ecological systems of foraging rabbits, the dynamics of how rabbits move around may not be fundamentally indeterministic, and may be perfectly predictable if, for example, the position, visual cues and neural states of the rabbits are perfectly known. Yet, we may choose to ignore such details, and to characterize the state of a rabbit in terms of position only. This is obviously an underdetermination, and multiple outcomes will be possible given the same position. In this way, coarse-graining and even ignoring certain variables can give rise to probabilistic causal relations (see Strevens 2003; Matthen 2009). For the

purposes of this chapter the precise nature of these probabilities need not concern us further, and we will treat them simply as given.

The notion of causal symmetry can be assigned different scopes, some more local, others more global:

Definition (localized to time and outcome)**.** *A causal network is* ***causally symmetric towards outcome o at time t*** *when the biases of any two states s and s' at time t towards o are equal: $P_s(o) = P_{s'}(o)$.*

This notion of symmetry is relevant for the question as to whether a particular instant in the past matters for a particular outcome. When the explanatory interest concerns the question whether *any* past state matters for a particular outcome, the following scope of symmetry is more appropriate:

Definition (localized to outcome)**.** *A causal network is* ***causally symmetric towards outcome o*** *when the biases of any two states s and s' towards an outcome o (at any time t) are equal: $P_s = P_{s'}$.*

This type of symmetry corresponds most closely to how path-dependence was formulated in the causal tree formulation, except that now an allowance is made for multiple possible initial states. Symmetry can also be localized to time alone:

Definition (localized to time)**.** *A causal network is* ***causally symmetric at time t*** *when the biases of any two states s and s' at time t (towards any outcome o) are equal: $P_s = P_{s'}$.*

Finally, a properly 'global' notion of symmetry can be formulated, as to predicate path-dependence about an explanation as a whole, not just an outcome: a network can be said to be **causally symmetric** when it is causally symmetric at every time t (or equivalently, towards any outcome o). This concept of global symmetry entails the three local notions of symmetry, and the most localized notion of symmetry is implied by the three others.

The transformation group associated with global symmetry is the group of permutations of the past states at any given time. Global symmetry arises when the conditional probability distribution over the outcomes remains invariant under permutation of the past states at any given time.

Figure 2.7 illustrates how these four scopes of symmetry can diverge. First, the network is not globally symmetric, since, for example, $P(o_1|s_4) = 1/2 \neq 0 = P(o_1|s_5)$. Thus, in order to explain why o_1 occurred, it is relevant that s_4 and not s_5 occurred. However, the network is symmetric with regards to some outcomes at some particular times. For example, at t_5 the network is symmetric towards o_2 as $P(o_2|s_3) = P(o_2|s_4) = P(o_2|s_5) = 1/2$. It does not matter what state the system is in at t_5 to explain why o_2 occurred. Similarly, the network is symmetric towards o_3 at t_3.

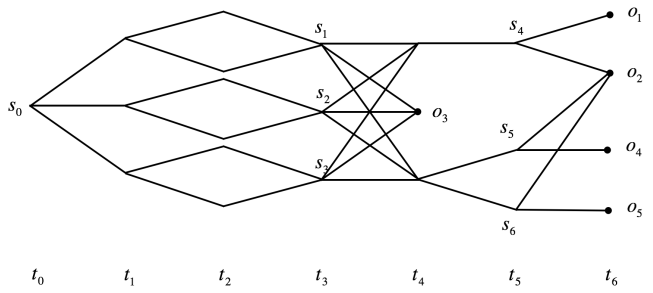

Figure 2.7: An illustration of how the different notions of symmetry come apart. Each state branches out in an equiprobable way.

Concerning the two other notions of symmetry, the network is symmetric at t_3, as the biases of s_1, s_2 and s_3 are equal towards *any* of the outcomes o_j. In an explanation of any outcome, it will not matter what state the system was in at t_3. Finally, the network is symmetric towards o_3. To explain why o_3 occurred, it will not be necessary to integrate *any* information about the past. Regardless of the path the system took, o_3 would have occurred with probability 1/3.

Deepening the parallel with spatial symmetries, causal symmetry can be given a geometric interpretation within a causal network. A network is symmetric at time t if every state at that instant t branches out to all descendant states in an identical way. Thus the branching pattern emitted by one state must be mirrored by all other possible states at that time. This basic pattern is represented in Figure 2.8, where the thickness of the lines is a measure for the probability of the different transitions. Some descendant states may be very improbable while others may be heavily biased; what matters is that the biases are symmetric across the differ-

ent initial states. At a symmetric instant in the network, the different states can be exchanged and permuted without the causal structure being affected.

This basic pattern of symmetry is both *maximally divergent* and *maximally convergent*. It is maximally divergent because each state branches out towards all possible descendant states; it is maximally convergent because each descendant state is converged upon by all possible predecessor states.

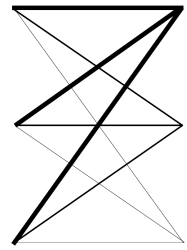

Figure 2.8: Fundamental pattern of symmetry.

Anticipating the next section, where symmetry is linked with path-independence, this fact suggests that path-dependence is to be sought between the extremes of perfect predictability and perfect unpredictability. Both the perfectly predictable network — where all paths converge onto a single outcome — and the perfectly unpredictable network — where all states diverge maximally — are ahistorical. Path-dependence requires some degree of unpredictability, but maximal unpredictability contingency implies path-independence. This is a concrete result that precludes any subsumption of unpredictability contingency under historicity (*e.g.* Beatty (2006)).

An additional effect of the basic pattern of symmetry is one of *erasing history*. In Figure 2.7, the network up until t_3 could be replaced by any arbitrary causal network without any difference being made to which outcome obtains. This effect is encapsulated in the following result:

Theorem 2. *If a causal network is symmetrical at t, it is also symmetric at all prior instants. The bias of any state towards a given outcome is shared throughout the states at any given time, and is preserved over time.*

Thus a sufficient condition for global symmetry is that only the last causal transition is symmetrical, *i.e.* each direct parent of the outcome states branches out to all outcomes. Note that, given such a symmetrical structure, none of the past states affects the outcome, and hence there is no history to erase, strictly speaking. History matters only to which intermediary states occur, and before the occurrence of the symmetrical pattern, it is possible to reconstruct the past. Once a symmetrical pattern occurs,

such reconstruction is impossible.[7]

A concrete example that could be represented by such a causal structure is mass extinctions. To the extent that one can idealize mass extinctions as the random selection of certain phenotypes (without regard to fitness), it is impossible to reconstruct the distribution of phenotypes *before* the mass extinction given the distribution *after* the extinction.[8] Even though non-symmetric processes may have dominated up until the point of the mass extinction, once the mass extinction has taken place, the effect of these processes on history is wiped out.

2.2 Symmetry Breaking

These different notions of symmetry are different ways in which the past does *not* matter, different ways in which the system is *in*dependent of the path taken. Path-dependence itself can be formulated as the *breaking* of symmetry, and thus has different scopes as well.

Definition (Path-dependence - symmetry formulation)**.** *A causal network is path-dependent relative to a certain scope if and only if the symmetry relative to that scope has been broken.*

In this way, a network may be globally path-dependent even though at certain times it may be path-independent, or even though certain outcomes may emerge in a path-independent way.

The attractor and causal tree formulations of path-dependence can be seen as special cases of this more general definition. If a causal network converges onto a global attractor, this means that any two states s and s' at any time t will lead to the outcome with probability 1: $P_s(o) = P'_s(o) = 1$. Conversely, if the outcome is not a global attractor, there is at least one possible state that is not in the basin of that outcome. In this case there are at least two states s or s' that have a different bias towards the outcome at some time t: the symmetry towards o is broken.

[7] Another implication is that while history may matter for the occurrence of some intermediary state, it is impossible for history to matter *for an outcome* at some time in the past but not ultimately (compare with Desjardins 2015).

[8] In this way, while mass extinctions introduce contingency into evolution (as famously emphasized by Gould 1989), to the extent that they make the reconstruction of the past more difficult, they actually remove some degree of historicity.

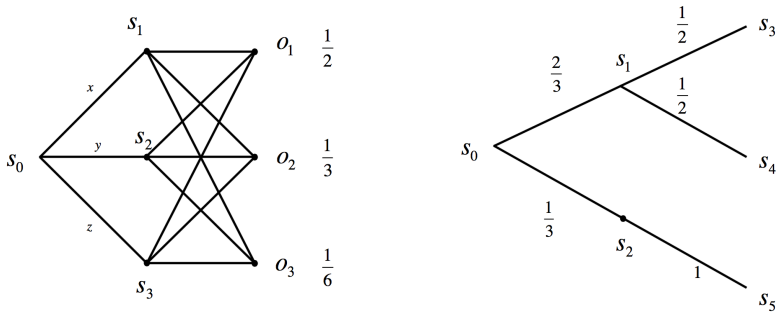

Figure 2.9: Unconditional probabilities of outcomes vs. path-dependence.

In the tree-framework, path-dependence was limited to comparing possible paths leading to one of a number of possible outcomes. The requirements of the causal tree formulation of path-dependence – there must be multiple possible outcomes (*i.e.* so that convergence can only be partial), and that paths towards some outcome affect the probability of the outcome – are captured within the negation of symmetry (localized to time and outcome). These requirements can be deduced by the condition that at least two states on different paths have a different bias towards the outcome.

The significance of this definition may be further illustrated by pointing out what it does *not* entail. First, it does *not* entail that no outcome is probabilistically privileged. Some outcomes may be more likely than others, and yet the network is symmetrical; all outcomes may be equiprobable, and the network path-dependent (Figure 2.9). The unconditional probability of an outcome is irrelevant; what matters is whether the conditional probabilities are equal or not.

A second orthogonal distinction is between path-dependence and the probabilities of the paths. The occurrence of an outcome may be path-independent, even though some paths may be heavily biased. For example, on the left side of Figure 2.9, there are three possible paths towards o_1. Even if the system may be much more likely to pass by s_1 than the other intermediary states (*e.g.* $x = 0.98$ and $y = z = 0.01$), $P(o_1|s_i) = 1/2$ for each intermediary state s_i. History does *not* matter: it makes no difference whether the system takes the s_1 path or the s_2 path, in each case o_1 will

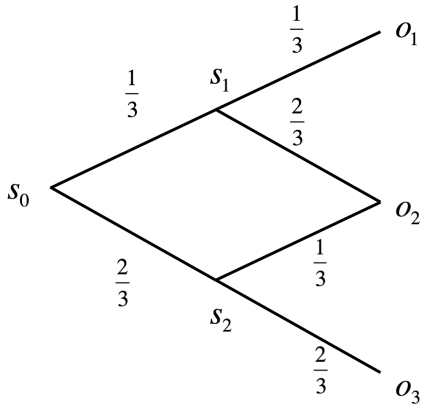

Figure 2.10: Symmetry towards o_2 is broken, even though all paths to o_2 equiprobable.

obtain with probability 1/2.

Thus, in a path-independent network it may be possible to reconstruct the past; conversely, retrodictability may be impossible in a path-dependent network. Such is the case in Figure 2.10, where the two possible paths towards o_2 are equiprobable, but yet where the symmetry is broken at the intermediate states since $P(o_2|s_1) = 2/3 \neq 1/3 = P(o_2|s_2)$.

There is no retrodictability here since, given that the system is in o_2, it is equiprobable that the system passed through s_1 as through s_2.[9] The relation between retrodictability and path-dependence will be taken up again in the final section, but since this result may seem puzzling here, one can illustrate it with an example. Say that s_1 represents 'financial crisis' and s_2 represents the avoidance of a financial crisis. The outcome state o_2 is a state of revolution. A financial crisis may be very improbable, but yet, once it occurs, revolution may be very likely. Conversely, a revolution may occur spontaneously with a very small likelihood. Even though these two paths may be equiprobable, if society actually were to undergo a financial crisis, any historian could integrate this information to explain the outcome.

[9]By Bayes' rule, $P(s_1|o_2) = \frac{P(o_2|s_1)P(s_1)}{P(o_2)} = \frac{2/3 \cdot 1/3}{4/9} = 1/2$.

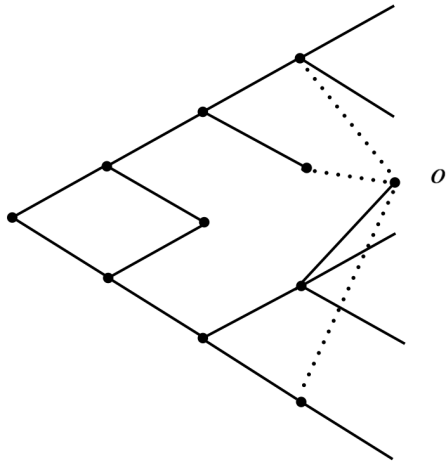

Figure 2.11: Weak global attractor: convergence and path-dependence.

2.3 Symmetry Preservation and Restoration

An additional advantage of the symmetry formulation is that it can distinguish between different scopes of path-dependence. Certain parts of a causal network may behave in a path-independent way, even though the network as a whole is path-dependent. The past may not matter in the causal explanation of a particular outcome, but yet may matter in the explanation of the set of outcomes. Or, the evolution of the system may be path-independent until a certain moment in time, after which the causal network becomes path-dependent. Path-dependence (localized to time) can emerge at a particular instant in the causal network.

Two combinations are of particular interest: cases where symmetry towards a particular outcome is preserved, despite global symmetry being broken, and cases where global symmetry is restored for a subset of the causal network. An example of the first is represented in Figure 2.11. Here the outcome o is a global attractor in the sense that all possible initial states can evolve towards o, and the occurrence of o is path-independent as all prior states are equally biased towards o. Yet global symmetry is broken in the network as a whole.

Such a state o can be termed a *weak global attractor*: a state that remains a possibility with a fixed probability regardless of the path the

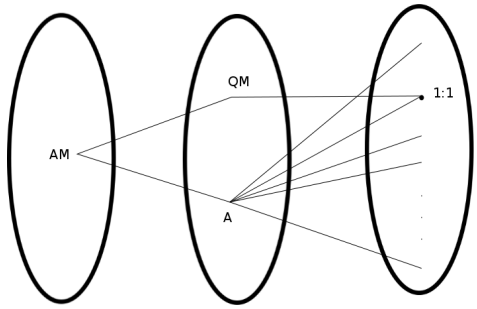

Figure 2.12: The ancestral monkey population (AM) branches into quadrupedal monkey (QM) and ape (A). The latter state has the capacity to evolve any limb ratio; the former can only keep the 1:1 ratio.

system takes. When a weak global attractor is present in a network, a local symmetry is preserved, even though the global symmetry may be broken.

The second case of particular interest concerns states that branch out towards all possible descendant states in an equiprobable way. Evolvability would be a concrete example of this causal structure.[10] For example, in most mammals, forelimb and hindlimb are locked by certain developmental constraints in a 1:1 ratio. A species can evolve longer hindlimbs only if the forelimbs grow by the same amount. However, in ancestral ape populations, a proper subset of quadrupedal monkeys, this constraint was relaxed, to allow for different possible ratios. A more formal representation would look something like figure 2.12.

Once the intermediary state A is realized, which outcome state (limb ratio) actually reached depends on the environment. In an extreme case, if absolutely no information about A's environment is available, all possible outcomes are to be modelled as equiprobable. This means that once A occurs, it no longer matters for the outcome what preceded that state. The causal network emanating from A constitutes a symmetrical causal tree. To the extent some outcomes can be privileged over others, symmetry is only restored to a certain degree (see next section). In either case, the state A can be thought of as a 'flexible' state: it partially restores

[10] The analysis given by Brown (2014) can be seen as dealing with this causal structure.

symmetry, limited to a subset of the whole causal network. Thus, while global symmetry once broken cannot ever be restored, global symmetry can be *locally* restored (to a certain degree).

3 Degree of Path-dependence

No account of path-dependence can be considered complete without giving some criterion of how history matters more in some processes than in others. We will focus only on how to quantify path-dependence according to how *informationally relevant* the past is for the outcomes. As already mentioned, a possible alternative way to measure path-dependence could be by quantifying how much an outcome changes if past states are changed. This would require the introduction of a separate metric (presumably dependent on explanatory interests) of what it means for outcomes to be close or distant, with associated problems (see section II). Instead, the focus will be on quantifying the degree of information given by the past in such a way that is consistent with the account of path-dependence presented thus also far.

3.1 Prediction and Retrodiction

In this approach, path-dependence is closely related to predictability and retrodictability in the following sense. An outcome is more predictable if a past state contains more information about which outcome will occur. Likewise, the past is retrodictable from the present if the outcome contains information about which causal path had been followed.

However, path-dependence precludes both perfect predictability and perfect unpredictability. Recall how a convergent network is perfectly predictable but path-independent, and a maximally divergent network is unpredictable but is also path-independent. In deciding then whether or not a network is path-dependent , it is thus irrelevant whether the outcome can or cannot be predicted from a past state.

The same point can be made about retrodictability. Thus it may be possible to know with fair certainty what causal path the system has followed, but for the network still to be symmetrical and hence path-independent. In Figure 2.9, we can know with fair certainty, given o_1,

that the system passed through s_1, even though passing through s_1 did not affect the probability of o_1. Retrodictability is possible despite path-independence towards o_1. Conversely, the outcome state may not contain any information about which causal path was followed, and yet the network can be path-dependent. This is the case in Figure 2.10, where both paths leading to o_2 are equiprobable, but where the choice between s_1 and s_2 affects the probability of the outcome.

The relation between predictability (retrodictability) and path-dependence can be made more precise by observing that the amount of information the past *contains* about the present is not relevant for path-dependence, but rather that the past *contributes* to predictability. Thus, in a network converging on a single outcome, the outcome is perfectly predictable regardless of whether the precise past state is known. However, knowing the past does not *contribute* any information not already contained by the structure of the causal network. Neither does it matter how much information the present contains about the past, but only how much the present affects retrodictability. In Figure 2.10, knowing which of the two intermediary states is reached helps to predict which of the three outcomes is likely to occur, whereas knowing which outcome occurred affects retrodictability.

One may wonder here if contribution to predictability and contribution to retrodictability are equivalent. If they were not equivalent, one would need to distinguish between two measures of path-dependence: a forward-oriented and a past-oriented measure. However, it is straightforward to show that they are equivalent.

Assume the past does not affect predictability, then the probability of an outcome conditional on an earlier state is simply the unconditional probability: $P(o_j|s) = P(o_j)$. Thus, in a network where the past does not contribute to predictability, the conditional probability of an outcome is equal to the unconditional probability. Similarly, it is the contribution of the present to the retrodictability of the past that matters. It does not matter when $P(s|o_j) = P(s)$ for every previous state s of a given outcome o_j. We would want to show that if $P(o_j|s) = P(o_j)$ for every outcome o_j and intermediate state s, then $P(s|o_j) = P(s)$ (and vice versa).

From Bayes' rule,

$$P(s|o_j) = \frac{P(o_j|s)P(s)}{P(o_j)}$$

and the desired result follows from the assumption that the past does not affect predictability. Thus it is impossible for the past to affect predictability without the present affecting retrodictability, and vice versa.

3.2 Mutual Information

Predictability is the lack of uncertainty of what the outcome state will be. Thus maximal unpredictability corresponds to a uniform probability distribution over the possible outcomes; maximal predictability assigns probability 1 to a single outcome and zero to the rest. In this way the **conditional entropy** of a set of outcome states O given a past state s,

$$H_s(O) = -\sum_o P_s(o) \log P_s(o) = -\sum_o P(o|s) \log P(o|s),$$

is a good measure for how predictable the outcomes seem from the perspective of past state s. It has a number of desirable properties: it is maximal for a uniform distribution, and zero when one of the outcomes is certain. A different conditional entropy, of O given a set of past states S at time t is obtained by taking the weighted average over the states in S:

$$H(O|S) = \sum_s P(s) H_s(O).$$

If it is known with certainty which state $s \in S$ occurred, then $H(O|S) = H_s(O)$.

The extent to which uncertainty is reduced by knowing which past states $s \in S$ occurred — the quantity, we have argued, relevant to path-dependence as symmetry breaking — is measured by the **mutual information** between the outcome states O and the set of past states S at some instant t:

$$I(O;S) = \sum_{o,s} p(o,s) \log \frac{p(o,s)}{p(o)p(s)} \tag{2.1}$$

Note that this formulation of mutual information is a measure of path-dependence localized to a particular instant in the causal network. Analogous measures can be formulated for the other notions of symmetry (both

local and global); however, the measure 2.1 is sufficient to extract the philosophically interesting properties.

Mutual information is consistent with the symmetry account of path-dependence in many different respects. First, mutual information is non-negative $I(O;S) \geq 0$, and zero if and only if the causal network is symmetric at $s \in S$. This can be seen as follows. If the network is symmetric at s, then for any given outcome state o and $s^* \in S$: $p(o|s) = p(o|s^*)$. From this and Theorem 2 can be deduced that these conditional probabilities are equal for all ancestor nodes, including any of the initial states s_0: $p(o|s) = p(o|s_0) = p(o)$. In this case $p(o,s) = p(o|s)p(s) = p(o)p(s)$ and hence

$$\begin{aligned} I(O;S) &= \sum_{o,s} p(o)p(s)\log 1 \\ &= 0 \end{aligned}$$

The mutual information is zero. The opposite also holds true: if mutual information is zero between O and S, then $p(o,s) = p(o)p(s)$ for every $s \in S$.[11] This implies symmetry.

Second, the claim that path-dependence is to be measured by information-contribution rather than information-content is underlined by the relation between mutual information and Shannon entropy.[12] Mutual information represents the information *gain* represented by an intermediate state:

$$I(O;S) = H(O) - H(O|S) \tag{2.2}$$

Thus, the degree of path-dependence is measured by the reduction in the uncertainty of the outcome states when information about later intermediary states S is integrated. Path-independence arises when there is no change in entropy content.

This suggests another way of viewing this aspect of path-dependence, in terms of the divergence of probability distributions. Mutual information can be expressed as the degree by which the unconditional $p(O)$ and the

[11] For the derivation, see *e.g.* Cover and Thomas 2006, Ch. 2.

[12] See Cover and Thomas 2006.

conditional distribution $p(O|S)$ *diverge*.[13] When there is no divergence, $p(O|S) = p(O)$ and the outcome states are independent of the past states S. Thus also in this respect, mutual information seems to be a natural operationalization of the symmetry formulation of path-dependence.

Third, mutual information is symmetric, *i.e.* $I(O;S) = I(S;O)$. This means that the present is relevant for the past in exactly the same way that the past is relevant for the present. This allows the previous arguments about the relation between path-dependence and predictability to be represented more formally. Here follows the case for predictability; identical reasoning can be applied to retrodictability (where $H_o(S)$ is the relevant measure for retrodictability). Perfect unpredictability means that the conditional entropy of the outcome states O is maximal, at any given set of past states S. This means that the unconditional entropy $H(O)$, which, in our framework, is the conditional entropy given the initial states S_0, is also maximal. Hence the mutual information $I(O;S)$ is zero, implying path-independence. Perfect predictability implies that the unconditional entropy $H_s(O)$ is zero at every past s; hence $I(O;S)$ is likewise zero.

This operationalization allows for information-theoretic analyses of path-dependence. Two interesting lines of inquiry for further research can be pointed to. A first concerns how mutual information changes as the grain of analysis changes. Thus, in the introduction we outlined how the path-dependence of a process depended on how both the initial states and the outcome states were described. The same process could be described as path-dependent and as path-independent. We already showed how fine-graining and coarse-graining had an effect on the convergence and divergence of a network; hence, one would expect the fine-graining of the outcomes to increase mutual information and thus path-dependence. With this in mind, we can conjecture that describing the outcome states at a more detailed grain of analysis increases the degree of path-dependence:

[13]The technical expression is that mutual information is the expectation, given S, of the Kullback-Leibler divergence between the distribution $p(O)$ and the conditional distribution $p(O|S)$:

$$I(O;S) = \mathbb{E}_S \left[D_{KL}(p(o|s)||p(o)) \right].$$

This is simply a quantitative expression of the how much the conditional probability distribution is expected to diverge from the unconditional distribution, 'from the perspective' of some time in the past.

In a given causal network (O, I, R), if $O = \{o_1, o_2, \ldots, o_n\}$ is fine-grained to $O' = \{o_{11}, \ldots, o_{1k_1}, o_{21}, \ldots, o_{2k_2}, \ldots, o_{n1}, \ldots, o_{nk_n}\}$, then $I(O'; S) \geq I(O; S)$.

The second line of inquiry would be to investigate how mutual information changes over time, and how it is affected by symmetry breaking.[14] For example, an interesting consequence of the nonnegativity of mutual information is that, through equation (2.2), the conditional entropy at some set of past states is never greater than the unconditional entropy: $H(O) \geq H(O|S)$. The entropy $H(O)$ can be thought of as the uncertainty on the distribution of outcome states without knowing anything about the past (*i.e.* the difficulty in reconstructing the outcome distribution). In this way the inequality means that 'information never hurts': knowing something about past states may turn out to be useless, but will never increase the uncertainty over the outcome states.

What is of interest is how the conditional entropy evolves over time $H(O|S)$. While a analysis in full generality is beyond the scope of this chapter, two simple cases can be mentioned. The first concerns the case where a network remains symmetric until some intermediate set of states S, after which the symmetry is broken. From (2) follows that the mutual information is zero at all past states S^* before S, and from (2.2), this means that $H(O) = H(O|S^*)$. Thus the conditional entropy remains constant until the breaking of the symmetry, after which it monotonically decreases. This is the same result, derived by different means, as in theorem 2.

A second simple case is when the network is a causal tree. Here $H(O) = H(O|S_0)$ (since there is only one initial state), and each branching even creates a sub-tree. Hence $H(O) = H(O|S_0) \geq H(O|S_1) \geq H(O|S_2) \ldots$, and conditional entropy monotonically decreases over time. In a branching tree, later states always contain more information about the outcome than the initial states do.

[14] See Sober and Steel (2011) for a related analysis of entropy change in Markov models. Since causal networks are Markovian, many of their results would also be applicable here.

Such a result may seem counterintuitive at first. Cannot a network start off with a bias towards some outcomes, and then evolve towards a uniform distribution, such as in Figure 2.13? Would this not increase the conditional entropy? The answer is that the network does not evolve towards a uniform distribution over *all* possible outcomes. The evolution towards uniformity is outweighed by the fact that any branch will have some inaccessible outcomes. Thus, while s_0 branches out to four different outcomes, s_1 and s_2 branch out to only two different outcomes. The entropy of four equiprobable outcomes is $\log 4$, whereas the entropy of two equiprobable outcomes is $\log 2$. In this case, $H(O|S) = \log 2$ and $H(O|S_0) = \frac{2}{3}\log 3 + \frac{1}{3}\log 6 > \log 2$. In this way, entropy also decreases here over time.

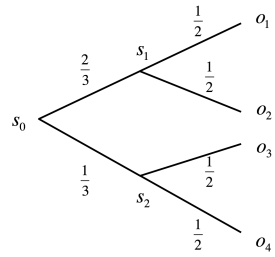

Figure 2.13: Decrease of entropy, despite uniformity.

Appendix

Theorem 1. *A coarse-graining of the explanandum makes an explanation increasingly convergent and a coarse-graining of the explanans makes an explanantion increasingly divergent.* s

Proof. We will prove it for an explanation that is purely parallel, thus neither convergent nor divergent. The generalization for a random explanation holds analogously.

Assume a deterministic explanation (O, I, f), so that f is a bijection $f : I \to O$. Define an equivalence relation $\sim$ on O such that $o_1 \sim o_2$ iff $o_1, o_2 \in A$ for some A (dependent on theoretical interests) with $\#A > 1$. Because f is a bijection there exists a uniquely defined $B \in I$ such that $f(B) = A$. Call B the 'basin' and A the 'attractor' of f on I.

Then O/A represents a coarse-graining of the explanandum and I/B a coarse-graining of the explanans. So define an associated function $R_c : I \to \#O/A : i \mapsto f(i)$ and relation $R_d \in I/B \times O = (f^{-1}(o), o)|o \in O$. Because f is a bijection, $\#I = \#O > \#O/A$ and $\#O = \#I > \#I/B$, and hence R_c will be a non-injective surjection, and R_d a non-function. Hence the number convergent structures has increased in explanation $(O/A, I, R_c)$, and the number of divergent structures has increased in $(O, I/B, R_d)$. □

Theorem 2. *Let (O, I, R) be symmetrical at some instant in time. Then (O, I, R) is symmetric at all prior instants.*

Proof. Assume (O, I, R) is symmetric at time t, corresponding to the set of intermediate states S. Let S' represent some earlier generation of states. From the local symmetry of (O, I, R) at S we can deduce that $P(o|s^*) = p \in [0, 1]$ for all $s^* \in S$.

Take a random predecessor state $s' \in S'$. Assume it branches out to a number of states $s^* \in S$. Then

$$\begin{aligned} P(o|s') &= \sum_{s^*} P(o|s^*)P(s^*|s') \\ &= p \sum_{s^*} P(s^*|s') \\ &= p \end{aligned}$$

since the sum of the probabilities of all paths leaving s' is 1. Thus the network is symmetric at S'.

This also means that the bias p towards outcome o is preserved as long as the network remains symmetric. □

Chapter 3

Ambiguity in Contingency Claims

In the previous chapter we saw how structures of path-dependence could be analyzed in causal networks; we will now outline how this could be applied to interpretations of evolutionary history. Interpretations of evolutionary history draw on hypotheses concerning causal mechanisms and possibility space in order to argue for a certain representations of evolutionary history, which then in turn ground claims about the contingency of evolutionary outcomes. The purpose of this chapter will be to show how these representations, which often remain only implicitly stated, can be explicitly formalized as causal networks. However, even then, causal networks may remain underdescribed, so that different claims about the contingency of outcomes may remain possible.

In the first two sections I will argue that the Gouldian and convergence-centric interpretations are ambiguous and that their contingency claims are insufficiently supported. This will require a certain reconstruction of Gould's and Conway Morris's arguments, since neither is explicit about their precise assumptions, or systematic about how precisely they understand the structure of evolutionary history (even though Conway Morris, in situating evolutionary history within hyperspaces, makes this reconstruction easier). I will try to carry out this reconstruction as neutrally as possible, and thus support it with textual evidence where relevant.

Once reconstructed with some more formal precision, I will then show

how both Gould's and Conway Morris's interpretations of evolutionary history do not support the claims they make about the contingency of evolutionary outcomes. Further assumptions would be needed, and these assumptions are controversial.

In a final section I will further systematize this critical analysis, and point to two general errors in drawing conclusions about the contingency of evolutionary outcomes. The first stems from using the probability of an outcome as a surrogate for the contingency of an outcome; the second from specifying the space of initial states with insufficient precision.

1 The Gouldian Interpretation

1.1 The Disparity Reduction Argument

Let us recapitulate the two informal arguments that we took in the introduction to underpin the Gouldian interpretation. The first – termed the 'disparity reduction' argument – is that certain developmental constraints that occurred early on in evolution affected all subsequent evolution; in particular, the selection of body plans without regard to adaptiveness during the mass extinctions of the Paleozoic constrained all subsequent evolutionary trajectories. In this way the initial large degree of disparity (number of distinct body plans defining a phylum) was irreversibly diminished, despite the increase in diversity of species later on in evolution.

The second is that there may be large-scale trends in measures such as complexity, but that this is simply the result of random exploration of biological possibilities, not of a causal bias towards higher complexity. Life started out from minimal complexity; hence it is inevitable that, given enough time, living organisms with higher complexity emerge. This trend can be expected even if evolutionary trajectories only randomly explore the space of possibilities. I call this the 'random walk' (from a left wall) argument.

Gould himself provides us with a phylogeny that illustrates the phenomenon of disparity reduction (Figure 3.1). He also illustrates the phenomenon with the following metaphorical characterization of evolutionary history ('Life'): "Life is a copiously branching bush, continually pruned by the grim reaper of extinction, not a ladder of predictable progress" (Gould

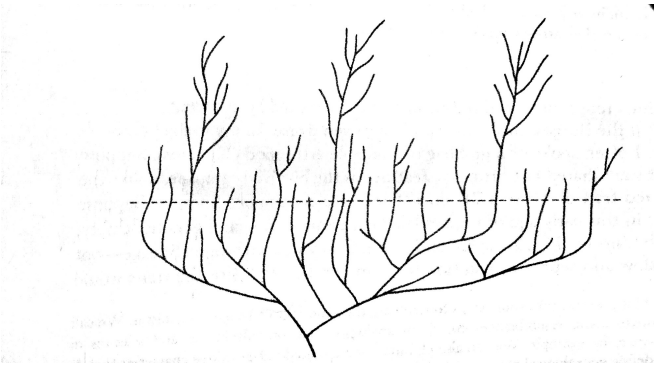

Figure 3.1: Disparity is at its maximum at the dashed line, and reduces thereafter due to a mass extinction event, or a sequence of such events. Diversity recovers to previous levels; disparity does not. (Reproduced with permission from Gould 1989.)

1989: 35).

What is crucial here is that the increase in disparity corresponds to an initial divergence in paths, mass extinctions cause most of these paths to stop at a certain point in time, and the few lineages that survive the mass extinctions go on to branch out once again (Figure 3.2).

However, by itself Figure 2 is insufficient to draw precise conclusions about the contingency of evolutionary outcomes. For that, we would need to know what would have happened if disparity reduction had occurred differently. What if different body plans had survived? Gould is very clear here: "any replay of the tape would lead evolution down a pathway radically different from the road actually taken" (1989: 51).

The only way this is possible is if the number of possible outcomes that living entities can conceivably attain is vastly larger than the number outcomes actually attained by evolution – a view so common as to be almost uncontroversial (Wright 1932; Kauffman 1993; Lenski 2004; Louis 2016; Smith 1970). The same view seems to underlie various statements made by Gould, for example, on the number of possible trajectories (1989: 50) or the long time it took for multicellularity to develop as evidence for "a vast realm of unrealized possibilities" (1989: 310).

In this way, we can place actual phylogeny (represented by Figure 3.1)

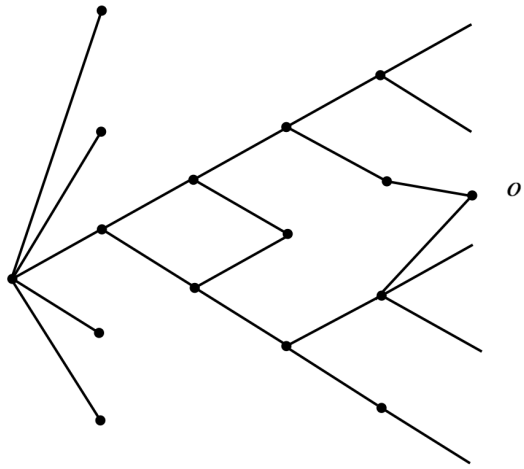

Figure 3.2: Disparity Reduction: an initial divergence is followed by subsequent divergences in a small fraction of lineages (with perhaps some limited convergences, such as the convergence towards o)

into the space of possible evolutionary paths (Figure 3.3), thus obtaining a representative causal network that condenses the disparity-reduction argument.

Note that Figure 3.3 is an *interpretation* of Figure 3.1 in the sense that it entails decisions about the degree of overlap between the possible divergences of evolution (shaded areas) with the actual divergence of paths (solid black area). If the space O_P was much smaller, one would expect an overlap to be more likely, but this is not necessary. As long as O_A is a proper subset of O_P, even a small possibility space would result in a similar causal network to that in Figure 3. Nonetheless, assuming there is no overlap is very plausible in the absence of convergent dynamics *and* if O_P is vastly larger than O_A – as is the case by multiple estimations, see below.

In this way, the step from Figure 3.1 to Figure 3.3 involves two assumptions: (1) biased branching towards O_A is unlikely since natural selection will not be able to overcome the developmental constraints obtained dur-

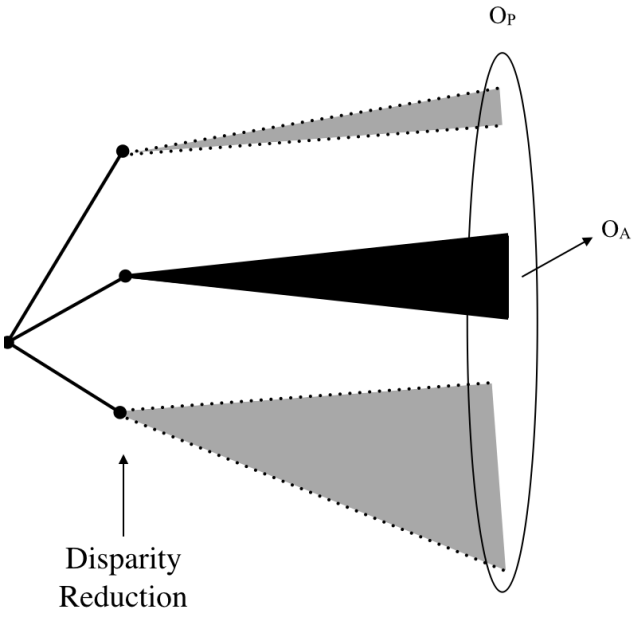

Figure 3.3: An initial divergence to maximum disparity is followed by a reduction in the number of actual lineages, followed by renewed divergence. Given the vast size of the set of possible outcomes (O_P), the set of actual outcomes (O_A) is likely to have no overlap with the sets of outcomes that could have evolved had the reduction in disparity occurred differently.

ing disparity reduction, and (2) the size of O_P is vastly larger than O_A, so that unbiased branching into is O_A also unlikely. Given these two assumptions, Figure 3.3 serves as an interpretation of Figure 3.1, in the sense that it places actual evolutionary history within a causal network of possible evolutionary histories. It is the latter that allows claims about the contingency of outcomes to be made.

More specifically, Figure 3.3 allows the nature of the contingency of outcomes to be read off from the symmetries present. At the initial state there is a symmetry towards all outcome states. This means that given the initial state of life, there is no favoured or privileged state. However, this symmetry is broken during disparity reduction, and the symmetry is

never restored. A bias towards O_A emerges. In this way, in the Gouldian interpretation, an initially uniform probability distribution over outcomes gives way to a probability distribution that is spiked in favor of O_A. Or, in information-theoretic terms, the mutual information between outcome states and states after reduction in disparity is *positive*.

Some nuances can be added to this paradigmatic causal network. First, this pattern of symmetry-breaking can be repeated again and again after the disparity-reduction event. This means that the intermediary state (or states) reached after disparity reduction could again function as initial state, from which a structurally similar evolutionary pattern could result. In other words, the pattern of symmetry-breaking could repeat itself over time in a self-similar way, at increasingly fine-grained levels of analysis.

Gould focuses on disparity reduction, since a body plan – determined by a number of factors, such as number of limbs, body segmentation, symmetry (see e.g. Gilbert 2000) – represents a major developmental constraint, and is difficult or impossible for selection to entirely overcome; however, there is a great variety of possible developmental constraints (and other types of constraint) which could induce further breaks in the symmetry. Such symmetry breaks could be less dramatic as a choice between body plans, and instead of making an area of possibility space inaccessible, they could merely make it less probable.

Second, Figure 3.3 does not preclude the possibility of convergent evolution. Within the solid black area, a number of evolutionary trajectories could independently arrive at a single outcome in O_A. However, such convergence is clearly limited, and does not diminish the contingency of the convergent outcome by much: evolutionary trajectories passing through the shaded areas could not converge on that outcome in O_A.

1.2 The Random Walk Argument

Gould's argument about large-scale trends can be formalized in different ways. The first would be to represent the diffusive process as a perfectly symmetrical network emanating from a single initial state, and steadily expanding to include the entire space of possibility. Such an expanding symmetrical network would trace all the possible paths a given lineage could take (simplified in Figure 3.4). A second, more coarse-grained representation of the range of expected evolutionary paths could be represented by some diverging cone over time[1] (Figure 3.6).

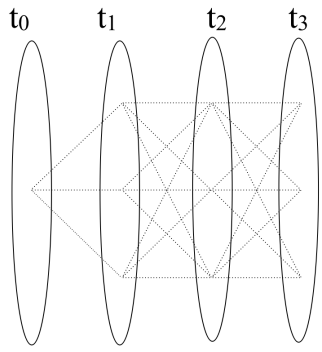

Figure 3.4: Symmetrical network diverging from single initial state.

In itself, the causal network implied by Gould's 'diffusion from left wall' argument is not very informative: a lot of information about the complexity of outcome states is not represented in the figures above. Hence, a third representation that captures some of the more interesting features would be to use the measure of complexity to induce an equivalence class on the outcome states, and hence another way of representing the evolution of complexity over time would be to count the number of lineages at a given quantity of complexity at a given time[2]. This gives a frequency distribution that both rises in height over time as well as becomes increasingly right-skewed (see introductory chapter).

What would change in Figure 3.5 if the tape of life were to be replayed

[1] This assumes that an initial state is more likely to branch out towards states 'closer' to the initial state. In a random walk – the simplest model of a diffusion process – the system is defined by a single variable, and the initial state can only branch out to its direct neighbours. Gould never goes into detail as to how precisely he understands the dynamics of a similar diffusive process in evolutionary history, but he assumes an 'expanding variation'. Representing this as an expanding cone can be seen as Gouldian in spirit.

[2] Note that we pass over how precisely 'complexity' should be defined – a contentious issue that we do not need go into here. Whatever the definition, the important thing the definition should do is to allow the different degrees of complexity to partition the space of possibility into different non-overlapping zones.

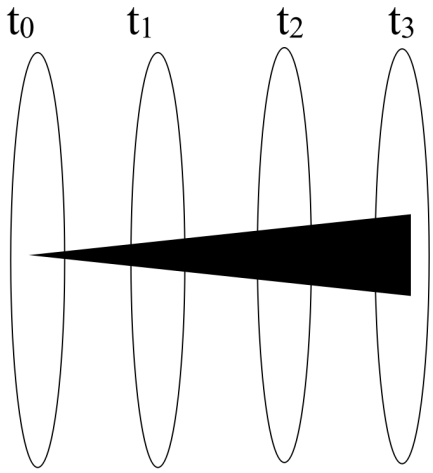

Figure 3.5: A diffusive process where variation steadily expands over time.

(from the same initial condition, but with different intermediate states)? In keeping with the disparity-reduction argument, assume symmetry-breaks lead to something like the following figure:
Note that this causal network is structurally identical to the one represented in Figure 3.3, and can be taken as the causal network that represents both of Gould's arguments.

Further, the divergence between the cones in Figure 3.6 masks the fact that degrees of complexity are similarly represented in the range of actual outcome states (O_A) as they would have been in some alternative range of outcome states (O_{ALT}). This is an illustration of how the contingency of an outcome is description-dependent: defined solely in terms of degree of complexity, the outcomes states are not contingent on past trajectories. In other words, there is a convergence of the totality of evolutionary trajectories onto an ensemble of trajectories where the degrees of complexity follow a right-skewed frequency distribution. This convergence is represented visually in the figure below (where $O_P^N = O_P \times O_P \times \cdots \times O_P$ represents the possibility space for N lineages).

Note, however, that according to Gould the convergence onto O_{RSD} is not due to any bias towards higher complexity in individual evolution-

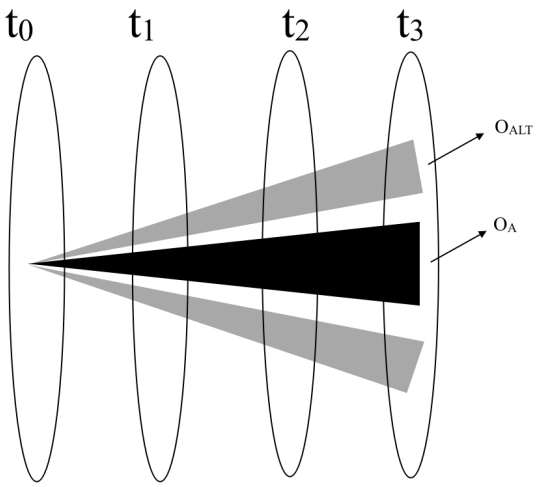

Figure 3.6: The range of actual outcome states (O_A) covers all degrees of complexity, even though a range of alternative outcome states (O_{ALT}) would do so as well.

ary trajectories. There is no evolutionary mechanism causing individual trajectories to evolve towards higher states of complexity[3]. Rather, the convergence onto O_{RSD} is simply due to the fact that it is the most *probable* outcome state for the ensemble of trajectories undergoing a diffusive process. Roughly, on the space of possible outcomes (O_P), O_{RSD} occupies the most volume, and hence any random path will likely end up O_P.[4]

[3]It is a result of "random motion from a simple beginning, not directed impetus toward inherently advantageous complexity" (Gould 1996: 173).

[4]This way of representing it draws on how diffusion is represented in microcanonical ensembles. The diffusion of gas molecules throughout a container occurs spontaneously because the number of ways in which gas molecules can be uniformly distributed across the whole container, is much greater than the number of ways in which gas molecules may be non-uniformly distributed, for example by being compressed into some subset of the container. In this way, the zone in phase space corresponding to uniform distribution of gas molecules occupies by far the most volume in phase space, and hence the system will evolve towards this zone with overwhelming probability.

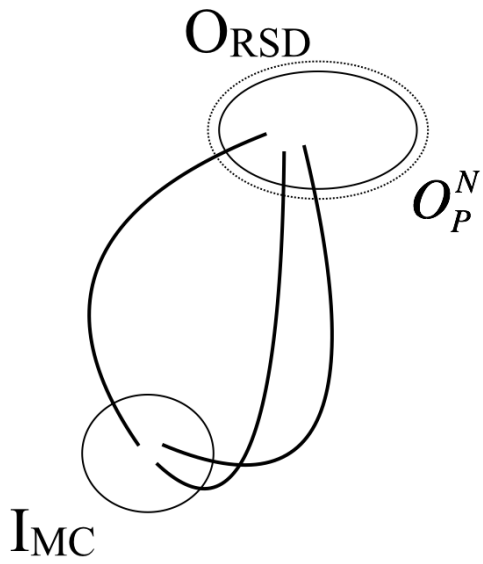

Figure 3.7: Gould's diffusion-from-a-left-wall argument implies a convergence from the set of initial states of minimal complexity (I_{MC}) onto the set of outcome states with right-skewed distributions over degrees of complexity (O_{RSD}).

Summary

In this way, the Gouldian interpretation of evolutionary history has two complementary elements: first, the continual breaking of symmetries by the emergence of 'frozen accidents' which constrain evolutionary trajectories, and second, an increase in complexity by diffusion from a left wall. As Gould notes, the inevitable increase in complexity allows for considerable contingency in the evolution of individual lineages:

> The right tail had to exist, but the actual composition of creatures on the tail is utterly unpredictable, partly random, and entirely contingent—not at all foreordained by the mechanisms of evolution. If we could replay the game of life again and again, always starting at the left wall and expanding thereafter in diversity, we would get a right tail almost every time,

> but the inhabitants of this region of greatest complexity would be wildly and unpredictably different in each rendition (Gould 1996:175).

In other words, the precise outcome of O_P is maximally contingent while the evolutionary outcome in (which represents the evolutionary outcome of all N trajectories) is not so contingent and converges onto O_{RSD}. For this reason, Figures 3.6 and 3.3 are perfectly compatible, and for purposes here, may be considered the representative causal network of the Gouldian interpretation.

We may further map the assumptions of the Gouldian interpretation according to how the initial and outcome state spaces, and the transition biases are specified. The only transition biases Gould integrates are biases arising from evolutionary constraints (such developmental constraints[5]) – biases arising from the mechanism of natural selection do not significantly affect the contingency of the outcome[6]. For this reason, the only dynamical principles underlying Gould's representation of evolutionary history are either (1) constraints which are themselves contingent occurrences in evolution, and, in the absence of constraints, (2) a random walk or diffusion.

Further, Gould has a double characterization of outcome space: (1) as morphological space (O_P), and (2) as the space of possible frequency distributions of lineages over degrees of complexity (O_P^N). In the disparity-reduction argument, he assumes that O_P is so vast that constraints imposed early on will likely not lead to any convergence in the attained set of outcomes. In the random walk argument, he assumes that O_{RSD} is so large that it occupies most of O_P^N. In this way, any given attained set of outcomes will most likely have a right-skewed distribution in degrees of complexity.

[5]See chapter 4 for a more in-depth discussion of the variety of evolutionary constraints.

[6]Even though Gould does not deny the possibility of convergent evolution, he plays down its significance for the contingency of evolutionary outcomes.

1.3 Ambiguity in Outcome Space

We will now argue how these assumptions are ambiguous, and in fact allow for different conclusions concerning contingency to be drawn. According to how details are filled in concerning outcome space – pertaining to the size of O_P, and to the size of O_{RSD} relative to O_P – we obtain contradictions, and claims about the contingency of outcome states (and contingency of a trend in complexity) are left insufficiently supported.

First, in arguing that the distribution of complexity for all species should get an increased skew over time, Gould reasons that "since space remains available away from the left wall and toward the direction of greater complexity, *new species occasionally wander into this previously unoccupied domain*" (Gould 1996: 171, my emphasis). However, it is not uncontroversial to assume that new species may occasionally wander into unoccupied domains of complexity. If the number of existing species is vastly smaller than the number of species that could exist at minimal complexity, it is possible that all of evolutionary history would take place entirely within the domain of minimal complexity.

Slightly simplifying the problem, assume that m_1 represents the maximal occupancy for the first degree of complexity, m_2 for the second, and so on. Hence, as the number of species N increases over time, once $N > m_1$ there will necessarily be a 'spill-over' effect, and the second degree of complexity will no longer be unoccupied. This shows that the increase of complexity is inevitable as the number of species increases over time.

However, what is the value of m_1? As an illustration, it is also a mathematical truth that, given enough time, a pot of boiling water will spontaneously freeze, as a thermodynamic ensemble explores all the possible microstates available to it. However, there is a catch: one will need to wait several times the expected lifetime of the universe for this to occur even once. Similarly, who is to say that m_1 is not so high that we could still be waiting – some 4 billion years after life began – for life to exit the bacterial mode? (Of course, then there would have been no observer to get impatient in the first place: see the discussion on observer selection effects.) Gould is entirely silent on this crucial issue. In fact, if m_1 turned out to be astronomically high, this would constitute good evidence for the hypothesis for a non-random exploration of the possibilities.

Analyzing the increase in complexity in terms of maximal occupancy

may capture the main idea but is somewhat simplistic. Gould is not claiming a spill-over into increased complexity, but a probabilistic increase in complexity. A particular zone of complexity need not be fully occupied before some lineage starts exploring the zone of higher complexity. However, even then, the volume that the minimal complexity zone occupies in possibility space is inversely related to the probability of a lineage ever exiting that zone.

As an example, consider protein folding, the process by which a protein changes the angle between amino acids in order to settle on a spatial configuration.[7] Such protein folding is necessary for proteins to carry out biological functions, and has a vast outcome space, with roughly 10^{150} possible spatial configurations for a protein of length of 150 (Louis 2016, Levinthal 1969). Louis notes that, at a sampling rate of 10^{13} tries per second, it would take 10^{120} times the age of the universe (10^{17} seconds) to sample all possibilities. Yet, in experiments, proteins settle on a particular spatial configuration within microseconds. He dubs this 'Levinthal's paradox', after its originator, and locates the origin of the paradox in a false assumption:

> Taking for granted that searching through the space of protein configurations is like looking for a needle in a haystack or like finding a hole on a very large flat golf course: each configuration is equally likely to be scrutinized. (Louis 2016: 112)

This argument assumes that certain life forms are 'adjacent' to complexity increases while others are no. This assumption is based on the thought that the dynamics of a true random walk should consist of a random permutation of the organismic structure of a lineage. This permutation is indifferent to complexity: some permutations will lead to complexity increases, others to decreases, and most permutations will presumably entail no change in complexity. Further, some structures need more permutations from others to achieve complexity increase; i.e., some structures

[7] In estimates of the number of possible organism configurations, protein spaces are often taken as surrogates, since their size is easier to estimate. Spaces of possible organism configurations are assumed to be vastly larger than these protein outcome spaces (see discussion between Salisbury 1969 and Smith 1970, but also Dennett 1996; Morris 2003; Louis 2016).

would need more random walk than others to increase in complexity. In this way, if the dynamics is a true random walk, some structures should be considered 'closer' to complexity increases than others. Conversely, if any possible structure may undergo a complexity increase, this suggests a dynamics other than a random permutation of organismic structure. In such a case, every state would have an equal probability of leading to a complexity increase, and we would obtain a weak attractor structure, where there is a convergent evolution onto states of increased complexity.

Another possible way out of this argument would be to deny that the volume of each complexity zone is vastly larger than the volume actually occupied by evolution, so that the probability of complexity increase is small but not negligible. However, since the degrees of complexity partition O_P, this means that the vast majority of 'complexity zones' remain unexplored, and that actual evolution remains limited to the lower complexity zones. However, since it is also possible that evolutionary histories start out in the minimal complexity zone, and remain constrained in a relatively small space (relative to O_P as a whole), this contradicts the argument that a divergence in developmental constraint early on precludes convergence later . In other words, outcomes would no longer be contingent as the disparity-reduction argument claims.

A second problem, compounding the first, is the observation of an *acceleration* in the maximum degree of complexity attained in evolution. First, note that there would seem to be *more* possible ways of organizing a more complex organism than there are of organizing an organism of lesser complexity. In terms of pigeon-holes, this means that $m_1 < m_2 < m_3 \ldots$ If the increase of N were to remain constant, one would expect a *deceleration* instead of an acceleration in the degree of complexity, as it takes longer to 'fill up' each container. Unless, of course, evolutionary history should not be understood simply as randomly exploring biological possibilities – for example, this could be taken as evidence that that there are in fact evolutionary mechanisms that bias lineages towards complexity increase[8].

One potential response is that the number of lineages accelerates at a greater rate than the various m_i. The acceleration of evolution is due to the

[8]Such a mechanism need not necessarily be based on natural selection; for example, Kauffman (1993) uses the vastness of possibility space to argue for self-organization. Louis argues for biases in how phenotypic variation appears.

great increase in the number of lineages, which speeds up the exploration of possibility space, despite the fact that the number of ways of realizing a degree of complexity increases with complexity.

However, this response underestimates the degree to which the size of possibility spaces may increase with complexity. The number of possible proteins compositions where 100 amino acids are picked from among a pool of 20 possible amino acids is $20_{100} \approx 10^{130}$, vastly greater than the number of protons in the observable universe (the Eddington number), which is estimated at around 10^{80}. The typical protein in the human body has 476 amino acids, and of this vast number only 20-50.000 different proteins are estimated to be present in the human body. At higher levels of complexity, estimating the number of possibilities quickly becomes intractable, more so because we often have no idea of the redundancy (the number of microconfigurations corresponding to a single macroconfiguration) – for example, presumably a large number of possibilities at the microlevel corresponds to nonfunctioning macroconfigurations.

A third incompatibility between disparity-reduction and random-walk arguments is that developmental constraints are problematic for conceptualizing the evolution of complexity as a random walk. Developmental constraints assumed by the first lead to 'generative entrenchment' (Schank and Wimsatt 1986) and irreversibility in evolution. This irreversibility – sometimes referred to as Dollo's law[9] – introduces an arrow of time different to the arrow implied by driven trends.

In this way, as a lineage embarks on certain paths rather than others, a large portion of the space of possible paths becomes effectively closed off. This consideration presents a problem for Gould's general view of evolution as an *effectively random* exploration of the different possibilities. As evolution unfolds it erects *barriers* to its own future evolution through developmental constraints[10]. In this way, Gould's argument about the effect

[9] An organization never exactly takes a former state again, even if it is placed under conditions of existence identical to those which it has crossed (Dollo 1905).

[10] Gould actually acknowledges this in a discussion of Seilacher's work, that organisms could have evolved body plans of "threads, ribbons, sheets, or pancakes so that no internal space lies very far from the outer surface" (1989: 311-312), with the result that internal complexity would never have been advantageous or even possible. However, in his later work (1996) he does not acknowledge the problems this poses for this random walk argument.

of body plans on evolutionary history can be turned against his view of a diffusion from a left wall: some body plans may have represented constraints against increases in complexity. This means that the increase in complexity turns out to be more contingent (contingent on certain developmental constraints not occurring) or less contingent (despite developmental constraints, complexity has nonetheless increased) than presented in Gould's view of a diffusion from a left wall.

2 Convergence-Centric Interpretation

2.1 Attractor Hyperlandscapes

The representative causal network to be associated with the convergence-centric interpretation is more difficult to represent visually, but Conway Morris gives us a more direct indication as to how precisely it is to be understood, so its reconstruction is more straightforward.

Conway Morris understands the possibility space within which evolutionary history unfolds as a vast 'hyperspace' with a very larger number of dimensions, and a very large range of possibilities. All evolutionary products, from the most simple to the most complex, can be situated in this hyperspace: "hyperdimensional boxes, based on the combinatorial immensity of the relevant variables that together encompass all the alternative possibilities (...), range from protein 'hyperspace' to intelligence and even societal 'hyperspace'" (2003: 308). From amino acid to human brains and societies, each outcome of biological evolution can be situated in this possibility space.

The dynamics within this possibility space is largely defined by convergent evolution – natural selection – which means that most regions of this hyperspace will never be accessed by evolutionary lineages because these regions are maladaptive. These regions – the "howling wildernesses of the maladaptive, the 99.9% recurring of biological space where things don't work" (2003: 309) – serve to severely reduce the 'effective' size of possibility space. By calculating the ways in which amino acids can be combined, the space of possible protein configurations is, by any estimation, vast (and similarly for other aspects of outcome space); however, most of these outcomes remain mere combinatorial possibilities, and are not accessible

given evolutionary dynamics.

In effect, Conway Morris is defining an *attractor hyperlandscape*, upon which evolutionary history unfolds[11]. As discussed in the previous chapter, attractor landscapes are limited for modeling dynamics of systems when the dimensionality of outcome space is very high. The importance of high dimensionality is that it tends to lead to large neutral networks on the landscape – so that evolution can occur without differences in the scalar (in our case, fitness) – and to a non-smooth topology, rugged peaks and holes – so that a relative difference in the scalar (fitness differences) has little to no predictive value for how evolution will occur.

In positing modeling evolutionary history as occurring on a hyperlandscape, Conway Morris is effectively stating that these potential difficulties associated with large dimensionality will not play a significant role, as least insofar as the contingency of evolutionary outcomes is concerned. Various examples of convergent evolution – legs, wings, hearing, seeing, smelling, cognitive functions (humans and dolphins) and even culture – are indicative of how most, if not all, evolutionary outcomes evolved.

In this way, the outcome space can be understood as mostly maladaptive, except for a limited number of attractor zones, which represent general biological traits, such as legs or seeing, but also "mammal-ness" and "ape-ness"[12]. Conway Morris's vision of evolutionary history could be thought of as a temporalized version of that of the pre-Darwinian Rational Morphologists (e.g. Goethe, Cuvier, Geoffroy St. Hillaire), who were preoccupied with finding the unchanging laws of organismic form. The space of biological possibility is surprisingly empty: there are only a few good ways of structuring life forms, and evolution hits on them again and again.

Finally, by limiting the outcome space in this way, most of these attrac-

[11] Conway Morris returns again repeatedly to the parallels with attractors, e.g. "Convergence occurs because of 'islands' of stability, analogous to 'attractors' in chaos theory" (2003: 127).

[12] "Life has a peculiar propensity to 'navigate' to rather precise solutions in response to adaptive challenges. I would suggest that one such solution is manifested in a biological property that we choose to call 'mammal-ness'. So, too, within this 'zone' there are more localized solutions, one of which is 'ape-ness'. (...) Although any history is necessarily unique, the resultant complex end form is not simply the contingent upshot of local and effectively random processes. On any other suitable planet there will I suggest be animals very like mammals, and mammals much like apes. Not identical, but similar, perhaps surprisingly similar." (2003: 308)

tor zones are inevitably realized in evolution given enough time – simply because the number of attractor zones is relatively small:

> First, what we regard as complex is usually inherent in simpler systems: the real and in part unanswered question in evolution is not novelty per se, but how it is that things are put together. Second, the number of evolutionary end-points is limited: by no means everything is possible. Third, what is possible has usually been arrived at multiple times, meaning that the emergence of the various biological properties is effectively inevitable. Finally, all this takes time. What was impossible billions of years ago becomes increasingly inevitable: evolution has trajectories (trends, if you prefer) and progress is not some noxious by-product of the terminally optimistic, but simply part of our reality (Morris 2003: xii-xiii).

If this were to be represented formally, one could discern two parts to the convergence-centric narrative. The first is that natural selection cuts down the space of possible outcome states to a limited number (Figure 3.8)

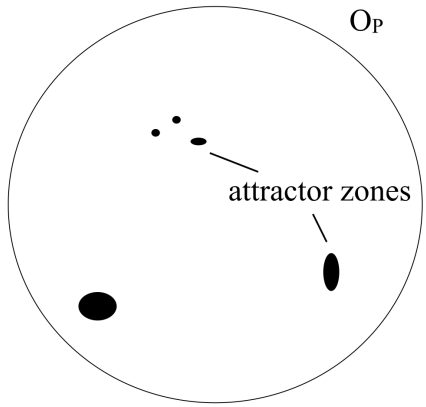

Figure 3.8: The reduction of the size of possibility space (O_P) by means of attractor zones.

The second is that, given a small number of outcome states, and given

enough time (on the order of billions of years[13]), all outcome states are likely to be visited. This can be represented either in analogy to Figure 3.6 (convergence onto the same range of outcome states), or Figure 3.7 (convergence onto the same type of distribution of convergent phenotypes over all lineages, i.e. a distribution where all the attractor zones are realized in at least one lineage): see Figure 3.9.

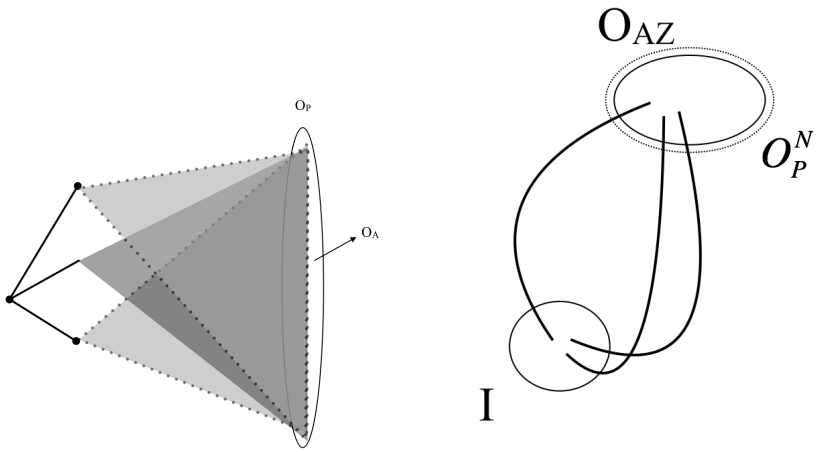

Figure 3.9: Alternative representations of the representative ensemble of possible trajectories implied by the convergence-centric interpretation. In the left figure, initial divergence is erased, since each evolutionary path leads to subsequent divergence in paths, and convergence to the attractor zones. In this way, the 'range' of attained outcomes (defined in terms of attractor zone) is the same. In the right figure, each initial condition converges to an effectively identical distribution of evolutionary outcomes (one where each attractor zone is realized in at least one lineage).

Because this representation of evolutionary trajectories is the one that Conway Morris draws on in order to ground his claims of inevitability (absence of contingency), we can take this to be the representative causal network of the convergence-centric narrative.

[13]Of course, given infinite time, it would not matter how many outcome states there were, as long as there were only finitely many.

2.2 Description-dependence in Outcomes?

Conway Morris defines evolutionary outcomes differently to Gould. The outcomes Conway Morris is interested in are outcomes such as 'seeing', 'mammal-ness', or 'ape-ness' – outcome-types that, to varying degrees[14], could be realized by a vast number of specific organisms (outcome-tokens, if you will). By contrast, Gould is interested in outcome-tokens when he claims that each replay of life's tape would lead to vastly different outcomes (an exception is when he is talking about distributions of complexity). At one point Conway Morris acknowledges that he and Gould are targeting different evolutionary outcomes (Morris 1998: 14), and thus talking past each other to a certain extent:

> What we are interested in [in contrast to Gould] is not the origin, destiny, or fate of a particular lineage, but the likelihood of the emergence of a particular property, say consciousness (Morris 1998: 14).

However, it is doubtful that this disparity alone can explain the difference between Gould's and Conway Morris's interpretations. Gould and Conway Morris were fully aware of this disparity, and yet their opposing interpretations remained unresolved.

Consider how Gould, following Darwin, acknowledges the distinction between "laws in the background" and "contingency in the details" (Gould 1989: 290). As examples of the first, Gould mentions the "laws of surfaces and volume", which constrain the shapes that larger organisms can take if they wish to preserve a similar surface to smaller organisms. Alternatively, there is bilateral symmetry in many organisms, which is a near-necessary consequence of cell division (1989: 289). Nonetheless, Gould states:

> Ultimately, the question of questions boils down to the placement of the boundary between predictability under invariant law and the multifarious possibilities of historical contingency. (. . .) I envision a boundary sitting so high that almost every interesting event of life's history falls into the realm of contingency (1989: 290).

[14] Presumably some types possess more possible relations than others (e.g., 'mammal-ness' vs. 'ape-ness').

In other words, Gould would also include outcome-types such as intelligence or 'mammal-ness' within the realm of contingency. The only outcomes he seems to exclude from the realm of contingency are those resulting from constraints imposed by mathematics (such as the relation between surface and volume as dimensions increase) or by physics (for example, the way in which the strength of gravity places a constraint on the minimal thickness of bones and other biomechanical properties).[15]

In this way, the different understandings of outcome space are already integrated within Gould's and Conway Morris's interpretations. Pointing to this type-token distinction among outcomes does not alter the conclusions that either of them reaches concerning the contingency of outcomes. Nonetheless, there is another ambiguity that goes unrecognized in the convergence-centric interpretation: an ambiguity in how the initial conditions are defined.

2.3 Description-dependence in Initial Conditions

A more important ambiguity in Conway Morris's interpretation of evolutionary history lies in how the set of initial conditions is determined.

Two alternative counterarguments to Conway Morris's view can be formulated. The first is that instances of convergent evolution do not imply that the outcome of convergent evolution is not contingent. Humans may be the result of convergent evolution in the way Conway Morris describes, but if the tape of life had been replayed in very different environments, very different instances of convergent evolution would have occurred. The reason for this is that natural selection produces *local* adaptations only: adaptations that are optimal in a given environment. If the environment changes, the same structures will no longer be adaptive. Hence natural selection can only select particular 'good moves' in design space when the environment is stable; when the environment changes, natural selection will change direction and what was previously a good move is no longer one entirely. This means that any structure that arises through natural selection remains contingent on a particular environment. If our planet were dark, eyes would never have evolved. One can only presume that the

[15]In this respect, Gould's notion of the 'realm of contingency' occupying most of biological possibility space is strongly reminiscent of Beatty's evolutionary contingency thesis (or vice versa: see next chapter).

evolution of a specific species of certain complexity, such as the *Homo sapiens*, would only be much more contingent on what environments occurred in the past. Thus, it does not matter how ubiquitous convergent evolution was on our planet, with our environments; it does nothing to show how the tape of life would have unfolded on a different planet.[16]

Turning to the second counterargument, let us assume that convergence in our biosphere was also a sign of convergence in all possible biospheres. In such a case convergent evolution would indeed establish inevitability. However, that would still not be very illuminating as to why a particular biological trait or organism evolved at all. Many different biological structures may be inevitable, so much so as to perhaps allow the formulation of a 'periodic table' of possible biological structures – but the number of inevitable biological structures would presumably be very large. As Schulze-Makuch et al. (2015) note,

> The last hundreds of millions of years of evolution on Earth provided us with a rich biodiversity of organisms, which explored a huge set of biochemical possibilities. Yet, our biosphere and the adaptations we observe are probably only a small subset of what is possible in biology. Life is intrinsically interwoven with its environment, so we can assume that during billions of years of natural history, many or nearly all of the biochemical possibilities were explored that are possible on a terrestrial planet with an average surface temperature of 15 C, 1 bar pressure and an oxygenated atmosphere (in the later part of Earth's natural history). (Schulze-Makuch et al. 2015: 1481)

Given enough time, life cycles through various combinations of these inevitable structures. This results in an interpretation of evolutionary history that is fundamentally identical to Gould's interpretation of large-scale trends: given enough time, a finite number of possibilities will be visited. In this way, the convergence-centric view of life only seeks to constrain the space of biological possibility, but it is possible that, even

[16] One way to counter this argument would be to show how certain structures are adaptive in a broad range of environments – some moves are 'good moves' everywhere, or almost everywhere. I call such structures 'general adaptations', and consider them in detail in chapter 6.

given the constraints of convergent evolution, the space of possibility is still vast so that evolutionary history is still effectively a random exploration of possibilities.

Conway Morris does not specify this wider space of possible outcomes, and assumes that the instances of convergent evolution on Earth establish inevitability and non-contingency. However, according to how the outcome space is construed, the convergence-centric interpretation may result in a view that is structurally identical to the Gouldian view.

3 Conclusion: Two Errors

3.1 Confusing Probability with Contingency

We can condense and systematize the discussion above by pointing to two common errors when reflecting about the role of contingency in evolutionary history. The first lies in using an estimate of the size of possibility space to draw conclusions about the *probability* of evolutionary outcomes, and then using these probability estimates as a basis for a claim about the *contingency* of evolutionary outcomes. Size of possibility space does not determine the probability of outcomes, and probability of outcome does not determine contingency.

Contingency is a concept that reflects the causal dynamics leading up to an outcome, whereas probability reflects the position of the outcome within possibility space. Probability and contingency are related, and both are description-dependent. Yet degree of probability cannot be used as a measure for the degree of contingency. We discussed this in chapter 2, but it is worth applying it to this context.

Two broad classes of cases may be distinguished where probability and contingency diverge. The first is when possibility space and *effective* possibility space are non-identical, for example, when large parts of possibility space are inaccessible, or if there is a bias towards certain areas in possibility space. Accessibility is a consequence of the dynamics in possibility space: an outcome state may be possible, and yet not realizable given the nature of the dynamics. The dynamics may also act to merely *bias* the system towards particular states (instead of to preclude certain areas). Certain outcomes – attractors – may be much more likely to occur than

others given certain initial conditions, or may even be more likely to occur given *any* initial condition.

Thus, estimates of the size of the space of possible genotypes provide little information about the relative contingency of various genotype-outcomes. Such arguments are variations on the following observation by Wright:

> Estimates of the total number of genes in the cells of higher organisms range from 1000 up (...) With 10 allelomorphs in each of 1000 loci, the number of possible combinations is 101000 which is a very large number. It has been estimated that the total number of electrons and protons in the whole visible universe is much less than 10100. (...) The population is thus confined to an infinitesimal portion of the field of possible gene combinations (Wright 1932: 356).

The 'error' may of course be easily avoided by reducing possibility space to effective possibility space, so that the probability of an outcome calculated as part of outcome space is identical to the probability calculated in terms of transition probabilities between states (see chapter 2). However, in order to do so, one must integrate information regarding evolutionary dynamics.

Some commentators take the size of possibility space as evidence that there must be mechanisms that help navigate possibility space (e.g. Dennett 1996 ('cranes' and search algorithms), Louis 2016 (bias in how variation is created)). However, for others, such as Gould, who hold that no evolutionary mechanisms introduce biases on a macroevolutionary scale, the size of possibility space is taken as a sign of the contingency of outcomes. The point here is not that either one of these possibilities is wrong; rather, it is that one needs certain assumptions about evolutionary dynamics to go from probability to contingency.

A second way in which contingency and probability diverge is when there is *redundancy* in outcome space: different outcomes that are treated as separate should actually be treated as 'equivalent'. This relates to the way in which Gould distinguishes between states (token outcomes) that Conway Morris would treat as 'equivalent' (outcome types). However, in his discussion of the large-scale trend in complexity, Gould does intro-

duce redundancy: outcomes are equivalent (and thus, in this sense, the difference between them 'redundant') if they possess the same degree of complexity.

Yet, Gould does not specify this redundancy with sufficient precision: *how many* outcomes are to be treated as equivalent given a fixed degree of complexity? He wants the redundancy to be not so big, so that an increase in complexity can happen spontaneously; yet, not too big so that the token outcomes are "utterly unpredictable and quite unrepeatable" (Gould 1989: 14). As argued in the previous section, these demands on redundancy are incompatible.

3.2 Confusing Convergence with Inevitability

If the first error involves lack of precision with regards to the outcome space, the second – interpreting convergent evolution as evidence for inevitability – reflects an ambiguity in the initial state space. Besides the fact that 'inevitability' means that the outcome occurs with probability 1, and thus confuses a question about contingency with a question about probability, there is also the mistake that evolution by natural selection does not imply that the outcome was symmetric over all possible paths.

The outcomes that occur through the mechanism of natural selection depend on the state of the environment in which evolution by natural selection occurs. Supplementary hypotheses would be needed to reduce the effective size of the initial state space, for example, hypotheses concerning the probability of certain environmental states. However, such hypotheses are not implied by the mechanism of natural selection. Relying on natural selection alone, one cannot escape the contingency of outcome states on initial states.

Conclusion

In this chapter we discussed some of the ways that description dependence influences interpretations of evolutionary history, and we took Gould's and Conway Morris's interpretations as test cases. Both Gould and Conway Morris describe outcome states differently; however, neither specifies his 'representative causal network' with sufficient precision. In Gould's interpretation there is ambiguity in the structure of outcome space; in Conway

Morris's interpretation there is ambiguity in the structure of initial state space. This chapter has attempted to sketch how these ambiguities entail that their claims about the contingency of outcome states are not entirely justified given their assumptions. Gould would need additional assumptions about the number of states associated with a given degree of complexity; Conway Morris would need additional assumptions about the probability of environmental states.

What is the appropriate level of precision? The level of precision is dictated by the hypothesis of what the key mechanisms are in evolutionary history. An interpretation of evolutionary history takes a stand on which mechanisms play a role in macroevolution; the challenge then is to specify initial conditions and outcome states in function of that assumption. Specifically, if natural selection is taken to be the key mechanism, then one needs to allow for the possibility that any environmental state may occur – not just those that are common on Earth.

In part II we will consider the causal basis of evolution by natural selection, and investigate both what 'representative causal networks' can be attached to this, and what claims can be made about the contingency of outcomes.

Part II

The Causal Basis for Interpreting Evolutionary History

Chapter 4

The Challenge of Causal Complexity

The first part of this dissertation examined how claims about the contingency of evolutionary outcomes should be analyzed in keeping with the fact that contingency claims vary according to the way in which the phenomena are described. Starting with the representative causal networks associated with interpretations of evolutionary history, we proposed in chapter 2 how contingency arises through symmetry breaks in causal networks. This also showed how speaking about the 'shape' of evolutionary history is a justified metaphor for evolutionary contingency: this shape is determined by convergences, divergences, and ultimately symmetries in the network of possible causal paths. The description-relative character of contingency has been insufficiently recognized by the proponents of various interpretations, and in the third chapter we argued that the claims made by Gould and Conway Morris about the nature of evolutionary history are insufficiently precise to fix their claims about the contingency of evolutionary outcomes. For example, their assumptions about the space of initial conditions or of outcomes are too ambiguous, and their claims about evolutionary contingency do not follow from their interpretation of evolutionary history.

The second part of this dissertation will dig deeper, and inquire into how interpretations of evolutionary history are grounded in biological theory. Biological theory shows up in the way in which interpretations general-

ize about the nature of causal transitions between evolutionary states. For example, some interpretations such as Conway Morris's hold that causal transitions are dominated by natural selection; others such as Gould's deny the dominance of natural selection, and claim that phylogenetic constraints are much more important. Conversely, when biological theory shifts, certain interpretations of evolutionary history may lose plausibility and relevance. Interpretations such as those of Henri Bergson (based on the principle of *élan vital*: Bergson (1911)) or Theodor Eimer (based on orthogenesis: Eimer (1898)), for example, can no longer be taken literally. Developments in biological theory necessitate an updating of interpretations of evolutionary history.

Hence we must look in more detail at the fundamental concepts in biological theory to examine how these concepts constrain possible interpretations of evolutionary history. Even then, biological theory is a sprawling patchwork of only partially overlapping concepts and theories. As a whole it may cover all known biological phenomena, but it is hardly a unified theory from which conclusions can be drawn for all possible evolutionary histories. Even the most 'fundamental' generalizations in biological theory, such as the Hardy-Weinberg 'law', are very limited in scope and do not apply to all of *actual* evolutionary history. Biological generalizations seem to be irreducibly contingent, and in one line of thought the contingency of the theory simply reflects the *complexity* involved in biological phenomena. One should not expect the dynamics of species or ecosystems to be as neatly describable as the dynamics of point masses.

Nonetheless, this poses a problem for interpretations of evolutionary history, since they do seek a unified representation of all possible evolutionary histories, despite the obvious complexity involved. Hence, interpretations of evolutionary history face what I call the challenge of *causal complexity*. The challenge goes roughly as follows: any generalization about biological phenomena seems to face *exceptions* because the generalization merely picks out one causal process (or a limited number of processes) in a complex totality of interrelated processes, so that there is almost always some process that was abstracted away from that may interrupt the target causal process, thus preventing the generalization from holding. More formally, the causal processes underlying any generalization 'all Fs are G' might be interrupted in many different ways, so that even if the

generalization is without exception in *actual* evolutionary history, there will invariably be a *possible* evolutionary history where it is only true for a subset of Fs, or where it is an empty generalization because Fs never evolved at all.

In the first section below I sketch this general background to the challenge of causal complexity. The challenge merely serves to orient the discussion; I regard it as insurmountable in its generality. Hence, in the subsequent sections, I will specify the scope of the discussion, and focus only on the fundamental concept of natural selection. In this way, I will not be examining how biological theory constrains interpretations of evolutionary history, but rather how biological theory constrains interpretations of *evolution by natural selection*. Here the challenge of causal complexity is more manageable, but by no means trivial.

In the second and third sections of this chapter I will consider what a relatively unified body of theory – the Modern Synthesis – implies about the contingency of directional trends that occur in evolution by natural selection.[1] This will set the stage for a reformulation of the challenge of causal complexity within the reduced scope of evolution by natural selection. The fifth and sixth chapters then are devoted to separate aspects of this challenge.

1 Contingency of Generalizations in Biology

1.1 The Evolutionary Contingency Thesis

An influential thesis advanced by John Beatty (and for which he explicitly cites Gould as an inspiration, see 1995: 45-46) states all generalizations about the living world are either (1) mathematical, physical, or chemical generalizations (or deductive consequences of such generalizations plus initial conditions), or (2) distinctively biological, in which case they are contingent. In other words, while the former may be true in all possible replays of life's tape, distinctively biological generalizations never are: they are merely contingently true. In this way, Beatty intends to capture and systematize the Darwinian/Gouldian distinction between 'laws in the

[1] In Chapter 6, we will examine the theoretical resources present in a successor, the Extended Synthesis.

background' versus the 'contingency in the details' by excluding the former from the 'distinctively biological'.

To support his thesis, Beatty shows how even generalizations that are central to the field, such as Mendel's first law[2] and the Hardy-Weinberg Law[3], do not achieve universal applicability, because they depend on the existence of certain entities or processes which themselves are contingent products of evolution. For example, gamete formation and sexual reproduction are all evolved traits which might not have occurred at all (Beatty 1995). In some possible evolutionary history, some alternative segregation ratio might have been the most common, so that Mendel's first law would have been largely false. In this way, the truth value of even canonical generalizations in biology is contingent.

Such examples of contingent generalization still seem subsumable under a D-N type model of explanation, except with more limited conditions of applicability. This is how Sober (1989) and Ereshefsky (1992) had earlier responded to the criticism that there are no noncontingent biological generalizations: for example, if the Hardy-Weinberg law is conditionalized upon its conditions of applicability, so that the logical structure of the law is not 'all Fs are G', but 'give C, all Fs are G', then this latter conditional is noncontingently true. In this way, part of Beatty's thesis can be accommodated within what is known as the ceteris paribus strategy, according to which many generalizations in special sciences may be considered exceptionless and confirmable as long as their scope is appropriately reduced (Fodor 1991; Lange 1993; Strevens 2012).

However, according to Beatty, the problem of contingency goes much further than this. To this end he articulates the difference between 'weakly' and 'highly' contingent generalizations. The weakly contingent generalizations are those whose truth is contingent on definite conditions. If these conditions are incorporated into the generalization, the original generalization is transformed into a non-contingent conditional that is true in every possible evolutionary history. In other words, weakly contingent generalizations are the type that can be turned into *ceteris paribus*-type

[2]Two alleles of a given locus have an equal probability of being represented in the gamete: a 50:50 'segregation ratio'.

[3]The gist of the Hardy-Weinberg law is thus: in the absence of evolutionary forces, genotypes of a population follow fixed frequency distributions determined by the relative frequencies of alleles. See the next chapter for a more detailed discussion.

generalizations.

By contrast, highly contingent generalizations cannot be transformed into non-contingent generalizations in this manner. No matter how the scope of such generalizations is reduced, they remain merely contingently true. One way in which such contingency can arise is through the multiple realizability of adaptive functions[4]. For example, it is impossible to generalize what type of reproductive contraptions orchids will evolve in order to fulfill the function of cross fertilization. Referring to Darwin's research into orchid reproduction, Beatty notes:

> Sometimes this part of the flower had been modified to entice or trap insects, sometimes another part had been modified to do the job. Even when the same parts had been modified to do that job, they did it in very different ways. (Beatty 1995: 58)

In other words, even given the same initial conditions (the same range of insects, the same need for cross fertilization, the same orchid morphology), very different outcomes are obtained.

In this way, Beatty is describing a phenomenon we analyzed in chapter 2: while a causal network driven by natural selection may converge on the same coarse-grained outcome (the biological function), the network becomes divergent when the outcome is fine-grained according to structure. At the fine-grained level of analysis, the divergence cannot be avoided by redefining the initial conditions (which is what the *ceteris paribus* strategy does), leading to an unavoidable contingency in outcomes in evolution by natural selection. This type of contingency precludes any exceptionless generalization about what precise outcome is to be expected.[5]

The phenomenon of the multiple realizability of adaptive functions has an even more far-reaching implication for evolutionary contingency than recognized by Beatty. The contingency in the realization of an adaptive function means that in further evolutionary history, evolutionary trajectories will diverge even further. Say that a function F can be realized by two different structures, S1 and S2. Yet these realizations may entail different

[4]Beatty uses the term 'functional equivalence'.

[5]Note that Beatty's notions of 'highly' and 'weakly' contingent parallel his distinction between unpredictability contingency and causal-dependence contingency.

developmental constraints on subsequent evolution. In this way, viewed on a larger temporal scale, this divergence in fine-grained structure may lead to divergence in coarse-grained functions later on.

1.2 Contingency of Outcomes and of Generalizations

Before we apply this to the types of generalizations relevant for interpretations of evolutionary history, it is important to analyze the term 'contingency' with some more precision, since there are two different uses of the term here[6]. The first sense refers to the *occurrence of an outcome*: if the outcome does not occur in some possible evolutionary history, it is contingent (to some degree). This is the type of contingency we primarily targeted in the first part, and is the type of contingency that interpretations of evolutionary history are concerned with. The second sense of contingency refers to the *truth of generalization* in possible evolutionary histories. A generalization may be thought of a proposition of the form of 'all Fs are G', where F refers to some evolutionary outcome (e.g. some individual, species or genus) and G refers to be some property instantiated by that outcome. Such a proposition may be true in all or only some possible evolutionary histories.

It is possible to integrate this second sense of contingency within the causal-network framework, and thus to show it as a special case of the first sense of contingency. Note that a generalization such as 'all Fs are G' can be used to induce an equivalence class on the set of Fs by dividing the set into two mutually exclusive subsets: the Fs that are G – call this $\mathcal{F}$ – and the Fs that are not G – call this $\mathcal{F}$. At a coarser grain of analysis, these equivalence classes can be represented as points, and thus as higher order evolutionary outcomes. In this way, the old analysis can be applied: a

[6]We pass over Beatty's rather idiosyncratic definition of the contingency of a generalization in his 1995 paper. There, a generalization is contingent if and only if it describes a contingent outcome of evolution. By 'outcome of evolution', Beatty means the 'rule-making capabilities of the agents of evolutionary change'. Further, by 'agents' he simply means evolutionary mechanisms, like "directed and random mutation, hybridization, natural and sexual selection, random drift, etc." (Beatty 1995: 47). By 'rules' he seems (somewhat circularly) to refer to generalizations, such as "humans are relatively hairless", "the Krebs cycle is present in all aerobic organisms" or Mendel's first law. My specification of generalizations as equivalence classes over evolutionary outcomes captures the essence of Beatty's definition.

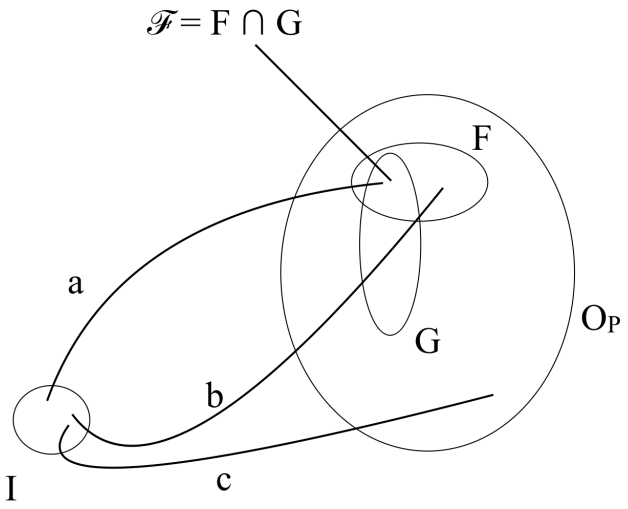

Figure 4.1: A rough analysis of the generalizations 'all Fs are G' in the causal-network framework. I is the set of initial states, three paths (a,b,c) illustrate three different evolutionary trajectories onto the set of possible outcomes (O_P). If a is representative of all paths, leading from I to $F \cap G$, the generalization 'all Fs are G' is necessarily true and applicable (given the initial state and evolutionary dynamics). If there are paths like b or c, then either some Fs are not G, or F might itself not be realized in evolution. In that case the generalization is either false (b), or inapplicable (c).

generalization 'all Fs are G' is contingent if its corresponding equivalence class $\mathcal{F}$ is a contingent outcome of evolution. This latter holds if there is some initial condition that leads to $\mathcal{F}$.

Analyzed in these terms, the question whether there are any non-contingent generalizations in biology is the same as the question whether there are any generalizations that are validly applicable across all possible evolutionary histories. The *ceteris paribus* strategy means that the scope of variation in initial conditions is reduced until there is convergence onto outcome $\mathcal{F}$.

1.3 Generalizations concerning Evolutionary History

What applies to biological generalizations applies a fortiori to generalizations that interpretations of evolutionary history tend to make. Two types of generalization can be distinguished.

In the first type, 'all Fs are G', F refers to a particular evolutionary history (set of evolutionary trajectories), and G refers to some property attributable to the evolutionary history as a whole. For example, 'evolution inevitably converges onto a limited number of adaptive outcomes', or 'evolution is invariably characterized by a tendency towards increase in body size', or 'evolution is utterly contingent' (which is the denial of non-contingent generalizations, and could be parsed as: for every F and G, 'all Fs are G' is contingent).

This type of generalization was the subject of part I. There we showed how G can be analyzed in terms of symmetries in causal networks, and by associating F with the representative causal network advanced by an interpretation of evolutionary history. Here we are more concerned with the second type.

The second type of generalization concerns the nature of causal transitions in evolutionary history. Specifically, in this type of generalization, F refers to a particular evolutionary state, and G refers to the causal process connecting F to its preceding and subsequent state along its evolutionary trajectory. Thus, for example, 'all significant transitions between states in evolutionary history are caused by natural selection', or 'outcomes in evolutionary history are caused by natural selection' would be a generalization that natural selection is a dominant mechanism in evolutionary history. Note that this second type of generalization – quantifying over possible evolutionary states – clearly informs the first type of generalization, which quantifies over possible evolutionary histories. Further, it is this second type of generalization that the challenge of causal complexity primarily targets.

1.4 Causal Complexity

Beatty's evolutionary contingency thesis *describes* a property of generalizations in biology, but does not *explain* why the generalizations should have these properties. It analyzes the structure of generalizations, but

without showing in detail how the generalizations reflect the subject matter of biology. He does give examples, such as the multiple realizability of adaptive functions, or the presence of 'random' and 'chance' mutations (Beatty 1995: 57-58). However, it is not clear why such seemingly indeterministic processes should necessarily occur in biology[7], or why the multiple realizability of functions should necessarily lead to contingency in evolutionary history.

A detailed investigation into the origin of contingency would lead us too far. For purposes here, it is sufficient to merely sketch an intuitive case why contingency should arise from the *complexity* of biological phenomena. In order to do so, we will review some arguments formulated by Smart (1963), Elgin (2006), and Kim (2005; 2012).

Smart originally compared special sciences such as biology and psychology to engineering:

> There are not any biological laws for the very same reason that there are not any laws of engineering. Writers who have tried to axiomatise biological and psychological theories seem to me to be barking up the same gum tree as would a man who tried to produce the first, second, and third laws of electronics, or of bridge building. (Smart 1963: 52.)

An engineer produces bridges, but bridges are complex constructions, and causally interact with the environment in a great number of ways that are relevant for their functioning. If asked how precisely bridges are to be structured, which materials in which quantities are to be used, and so on, the best answer an engineer would be able to provide is, 'it depends'. A bridge causally interacts in so many ways with its environment (type of soil underneath, the type of entity that will cross the bridge, the height and length of the bridge, etc.), that the microstructure realizing the bridge will be very sensitive to context, so much so that it is not possible to find any exceptionless generalization about the precise microstructure of bridges (beyond generic features).

[7]Some have claimed that the probabilistic character of genetic mutation is ultimately a manifestation of quantum indeterminism (Brandon and Carson 1996). This does not seem plausible, but even if this were true, it still does not explain why quantum indeterminism should be manifested at such large spatial scales (whereas quantum indeterminism plays a negligible role in the vast majority of macroscopic phenomena).

This insight has been restated by Elgin (2006:123), who drawing on Cartwright (1983; 1995) has argued that the superposition of the multitude of causal processes describing the behaviour of the parts makes it difficult to simply treat the behaviour of the whole as a simple addition. Consider the superposition of two *ceteris paribus* generalizations, each describing an aspect of a complex system: (i) *ceteris paribus*, adding salt raises the cooking temperature of water, and (ii) *ceteris paribus*, increased altitude lowers the cooking temperature of water. What if salt is added to a pot of water and it is brought to a higher altitude? The *ceteris paribus* generalizations by themselves do not provide sufficient information to give any answer to that question.

Thus, Elgin argues, a general argument could be made that, since (1) each of the parts of a complex kind causally interacts with the environment, (2) the global behaviour of the complex whole is the superposition of the behaviours of the parts, and (3) many of the causal processes defining part-environment interaction make a difference for the global behaviour, then the global behaviour is not simply an addition of the behaviours of the parts. One needs not only to weigh the relative importance of each part-environment causal interaction, but also to take into consideration that parts may interact among each other as well.

A second argument (Kim 2005, 2012) against the possibility of exceptionless generalizations in special science draws not on the complexity in causal interaction with the environment, but on the complexity of constitution. Kim assumes a connection between complexity and the variability between the tokens of a special-science kind. This variability, or 'idiosyncrasy', ensures that strict, exceptionless laws are unavailable.

Going forward, it is important to distinguish between two different senses of complexity. The first is *compositional complexity*: biological entities are not point-particles, but are composed of sub-entities. The second is *dynamic complexity*: this complexity is a measure of the number of independent variables needed to describe how a biological entity evolves over time. It is the latter type of complexity we primarily have in mind when speaking of 'causal complexity'.

The challenge of causal complexity can be thus formulated:

> Any generalization about the causal nature of transitions in evolutionary history is necessarily contingent, because there is

a large number of independent causal processes impacting each state in evolutionary history, ranging from external processes (mass extinctions, climate change, invasion of a predator, etc.) that either change selective pressures or lead to 'indiscriminate selection' (i.e. drift), to internal processes (developmental constraints, mutation) that alter the phenotypic variation in a population available for natural selection.

Conclusion

As suggested in the introduction, it is likely impossible to entirely overcome the challenge of causal complexity. In particular, it is doubtful that any generalization about the shape of evolutionary history – for example, that complexity tends to increase over evolutionary history as a whole – should not be subject to counterexample. This is why interpretations of evolutionary history often draw on certain fundamental concepts in biological theory, such as natural selection or developmental constraints. In this respect interpretations are very elaborate *ceteris paribus* generalizations: by focusing only on the consequences of these fundamental concepts for evolutionary history, they deliberately abstract away from and simplify the mass of causal details. As empirical hypotheses, they will undoubtedly succumb to exceptions; however, their deliberate abstraction away from causal detail does increase their explanatory value, because the abstraction allows for the shape of evolutionary history to become discernible.

Nonetheless, this abstraction and simplification is not an arbitrary process, and is severely constrained by biological theory. As an extreme example, it would be difficult to construct an interpretation of evolutionary history that claimed that the large-scale features of history can be explained by lightning bolts striking organisms at random (even though this presumably would not be logically inconsistent with biological theory).

The rest of this chapter is dedicated to an overview of a unified and very influential body of theoretical work: the Modern Synthesis. The Modern Synthesis makes key claims about how the space of biological possibility is to be understood (primarily in terms of gene frequencies), and about which mechanisms play a dominant in driving transitions between states (primarily mutation and natural selection). In this way, the Modern Synthesis contains a number of key tenets that clearly influence how

evolutionary history can plausibly be interpreted.

2 Elements of the Modern Synthesis

To understand why the different tenets of the Modern Synthesis are not simply arbitrary stipulations, it is instructive to consider the historical genesis of the Modern Synthesis, and the way in which these tenets were painstakingly arrived at as solutions for important problems. In the following we will give a very brief historical overview.

Darwin and Wallace were the first to propose the theory of natural selection, which however, initially faced a number of objections initially. A particularly important objection was the 'Swamping Argument', formulated in 1867 by Fleeming Jenkin in a review of On the Origin of Species. According to Jenkin, even if a favourable trait were to arise in a population, it would be quickly swamped through breeding between the fitter individual and other individuals in the population. There could be no novelty in evolution, and further, one would expect organisms in natural populations to quickly become uniform.

To properly understand the Swamping Argument, one needs to know that it relies on the hypothesis of blending inheritance. According to this hypothesis, the traits of the offspring are the average of the traits of the parents: for example, a large and a small parent will together produce a medium-sized offspring. Darwin worked with a particular hypothesis of blending inheritance that he termed 'Pangenesis'. Pangenesis is the theory that every cell in the parent body produces 'gemmules' or 'pangenes', which contain the blueprint for how such a cell should develop. These gemmules subsequently concentrate in the reproductive organs, from where they are then passed on to the next generation (Darwin 1868). These gemmules from both parents would then combine in the offspring to produce the traits of the offspring, even though they could also be passed on in a 'dormant' state, and be activated only a number of generations later (see Geison 1969). Further, because such gemmules could be produced by the cells at any time during development, the offspring would also inherit acquired characteristics[8].

[8]Hence Darwin was actually committed here to Lamarckian inheritance: the inheritance of used traits, and disinheritance of disused traits. Where he disagreed with

In short, Darwin worked with a rather detailed (but ultimately false) theory of inheritance, and this led to the Swamping Argument. However, at the time it was not known that the fault lay with the theory of inheritance, and not with the theory of natural selection. The argument was seen as a decisive blow to the theory of natural selection, and contributed to what Julian Huxley termed the 'the Eclipse of Darwinism' (Huxley 1942) – referring to the period between the 1880s and 1930s when only a minority of biologists believed natural selection was a major evolutionary mechanism.

Ironically, even before this controversy started, in the 1850s and 1860s Gregor Mendel had formulated an alternative theory of *particulate inheritance*, which claimed that traits did not develop from a fusion of parental traits, but from discrete 'factors' inherited from parents. His most famous experiment demonstrating this involved breeding two pure-line varieties of *Pisum* (pea) with regard to binary characteristics, such as the form of the seed (round or wrinkled) or the colour of the cotyledon or of the inside of the pea (yellow or green). Mendel bred the two pure-line varieties, resulting in a hybrid population of pea. This result was perfectly expected under blending inheritance; however, what was surprising was what happened when these hybrids were interbred. According to the blending inheritance the offspring of hybrids would again be hybrids. Instead, Mendel found *renewed* variance, following fixed ratios. These fixed ratios could be easily explained by representing the breeding as crossings between different 'factors', whence the diagram in Figure 4.2.

The distribution of traits in the third generation is also known as the Hardy-Weinberg equilibrium (it was not explicitly formulated by Mendel, but independently by G. H. Hardy and Wilhelm Weinberg in 1908). Note that such an equilibrium would be impossible under blending inheritance: instead, the equilibrium distribution would be uniformity.

Nonetheless, Mendel's results were almost entirely ignored. Cultural barriers may have been partly to blame; however, an important reason was that 'Mendelism', as particulate inheritance became known, seemed unable to account for *quantitative* characters (traits that vary continuously, such as height or hair colour). The continuity of the phenotype seemed at

Lamarck's view of evolution was in the role of inheritance for the transmutation of species: Darwin believed natural selection to be a much more powerful cause than inheritance of acquired characteristics.

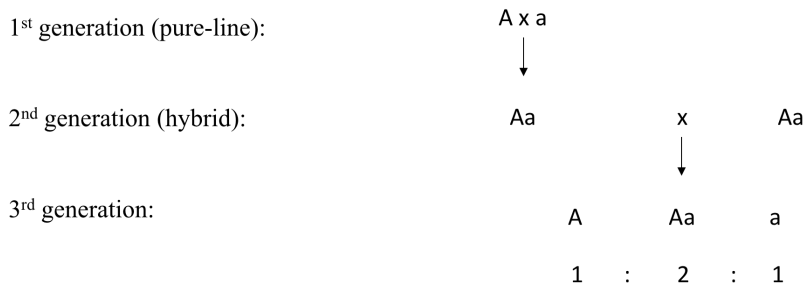

Figure 4.2: The Mendelian ratios.

odds with the discreteness of the 'factors' underlying them. Hence, while particular inheritance worked very well to explain the shape and colour of pea pods, it was seen as a curiosity instead of a general principle underlying all inheritance.

This problem was overcome by Ronald Fisher, one of the first architects of the Modern Synthesis, who showed in a 1918 paper that when multiple factors were responsible for a single trait (by contrast, Mendel considered only traits controlled by a single factor), and when environmental variations are taken into consideration, the result is a bell curve-like distribution of phenotypic traits. Discrete factors can lead to continuous traits. In a later paper Fisher also showed how Mendelism can address the Swamping Argument. He showed how under particulate inheritance, to maintain the same level of variance in a population, new variations need to occur much less frequently with particulate inheritance than with blending inheritance. In this way the work of Fisher both overcame an important obstacle for the acceptance of Mendelism, and revealed its great advantage over blending inheritance by showing how it resolved the Swamping Argument.

Many advances were made by applying (and in Fisher's case, inventing) statistical methods to Mendelism and natural selection. An important contribution concerned how novel species could arise without presupposing sudden large changes in phenotype (saltationism). A a model for this was outlined by Sewall Wright in 1932, which held that species arose when novel variations avoided extinction through drift by becoming isolated from the larger population. In this way populations could, by means of drift and

reproductive isolation, traverse a valley on what Wright called the 'adaptive landscape' and reach a different adaptive peak – which represented, by assumption, a novel species.

A final issue, orthogonal to particulate vs. blending inheritance, concerned whether acquired characteristics could be inherited or not (Lamarckian inheritance). Indeed, one of the first scientists who rediscovered Mendel's work following a decades-long neglect, Hugo de Vries, used the ideas to defend a particulate version of Pangenesis: the pangenes arise from every cell in the body, but instead of fusing, as Darwin had envisaged, they combine in Mendelian patterns.

It was the work of August Weissman which established the separation of genetic material from somatic material, preventing the latter from playing any role in inheritance, and thus precluding the inheritance of acquired characteristics. Famously, Weissman chopped off the tail of mice for five generations, and not once was a mouse born without a tail. Even though this experiment strictly speaking does not refute sophisticated forms of Lamarckism (which claimed only the disinheritance of *disused* characters, not of externally manipulated ones), Weissman did introduce an influential dualistic understanding of life: between genes (or 'germ-plasm' as Weissman called it) that exist in perpetuity, handed down through generations, and the organisms that house them.

In this way, the modern concept of the gene relies both on the Weissman barrier between soma and germline, as well as on the idea of particulate inheritance. Another result of these separate advances was a theory that went on to become a new science – population genetics – that showed how natural selection acting on genetic variation in a population is sufficient to drive evolution through gradual steps. The new theory of population genetics made other theories that posited unknown sources of mutation superfluous – such as saltationism (evolution occurs in big jumps) or orthogenesis (variations tend to occur in certain lines of development).

The synthesis of Mendelism with Darwinism, together with the Weissmanian genetic programme, is what is commonly known as the Modern Synthesis. There is some disagreement among biologists as to what precisely should be included and what excluded. Often the discovery of DNA by Watson and Crick and the 'central dogma' of molecular biology are included. This dogma holds that DNA determines phenotype through tran-

scription to RNA and to proteins, and that this arrow of determination cannot be reversed. In this respect the dogma represents a restatement of the Weissman barrier with more causal detail.

There is also disagreement as to how much of the Extended Synthesis is not already contained with the Modern Synthesis. For example, Laland, Wray et al. (Laland et al. 2014) argue that phenomena such as plasticity, niche construction and developmental bias (all to be discussed later on) can be reduced to the fundamental processes explicitly recognized by population genetics: natural selection, drift, mutation, recombination and gene flow.

In the following, we will identify the Modern Synthesis with the following five tenets (from Futuyma's textbook):[9]

> The major tenets of the evolutionary synthesis, then, were (1) that populations contain genetic variation that arises by random (i.e., not adaptively directed) mutation and recombination; (2) that populations evolve by changes in gene frequency brought about by random genetic drift, gene flow, and especially natural selection; (3) that most adaptive genetic variants have individually slight phenotypic effects so that phenotypic changes are gradual (although some alleles with discrete effects may be advantageous, as in certain color polymorphisms); (4) that diversification comes about by speciation, which normally entails the gradual evolution of reproductive isolation among populations; (5) and that these processes, continued for sufficiently long, give rise to changes of such great magnitude as to warrant the designation of higher taxonomic levels (genera, families, and so forth). (Futuyma 1998: 12; quoted in Pigliucci and Muller 2010)

Especially important for our purposes are tenets (1), (2), (4) and (5), which express a hypothesis of what mechanisms are dominant in evolution, and how this mechanism shapes evolutionary histories.

[9] Wray and collaboraters might disagree with identifying the Modern Synthesis with its historical form, see for example: "What Laland and colleagues term the standard evolutionary theory is a caricature that views the field as static and monolithic." (Laland et al. 2014: 163). The issue as to what precisely the Modern Synthesis and Extended Synthesis entail is a controversy I do not need to take sides on.

3 Contingent Directionality

Since the Modern Synthesis entails claims about how states of evolutionary history are to be represented (gene lineages), and what mechanisms dominate the transitions between states, it is a significant theory that clearly puts constraints on interpretations of evolutionary history. Both the Gouldian and convergence-centric interpretations are compatible with the Modern Synthesis (even though Gould emphasizes developmental constraints more than the Modern Synthesis typically does). By contrast, historical interpretations of evolutionary history, such as interpretations based on vitalistic principles (e.g. Bergson) or orthogenesis (e.g. Eimer), are, if not incompatible, at least not straightforwardly compatible.

3.1 Trends Driven by Natural Selection

The interpretation that fits perhaps most naturally with the Modern Synthesis is Dawkins's gene-centered view of evolution. According to Dawkins, evolutionary history is a story of competition between gene lineages: genotypes (or 'replicators') that engage in an evolutionary arms race in order to replicate as often as possible. Phenotypes (or 'interactors') are produced as weapons in this arms race, and the result is a view of evolutionary history as one of continual competition, by various means, between gene lineages (Dawkins 1982).

This permanent arms race can generate directional trends by causing predators and prey, or parasites and hosts, to coevolve. Yet, even if such trends are instantiated, the evolutionary outcomes they result in are contingent. As Dawkins states in *The Ancestor's Tale*, focused in particular on the outcome of the human species, but applicable to outcomes more generally:

> It makes no more sense (and no less) to aim our historical narrative towards *Homo sapiens* than towards any other modern species — *Octopus vulgaris*, say, or *Panthera leo* or *Sequoia sempervirens* (Dawkins 2004: 6).

It makes no sense to privilege one outcome over another, because all are equally contingent on selection pressures. Thus, while natural selection

may give rise to directional trends by predator-prey or host-parasite co-evolutionary dynamics, the actual outcome is contingent on the types of environment that arise.

In general, the Modern Synthesis allows for two types of directional trend to be realized in evolutionary history, where both trends are contingent occurrences. The first type of directional pattern is caused by natural selection. Natural selection occurs when individuals with different phenotypes reproduce at different expected rates in a population. These differential reproduction rates are typically caused by certain features in the organism-environment interaction. For example, if a moth population is placed in an environment of soot-covered trees, the darker moths will be less likely to be spotted by predators, and hence will be able, on average, to produce more offspring than the lighter moths. In this way, after some generations, the moth population will have evolved to contain only dark moths, which are camouflaged and in this sense 'adapted' to their surroundings.

Natural selection is often represented as an optimizing process, driving a population to an equilibrium state. However, this common picture is controversial. In chapter 5 we will discuss how those skeptical of the causal nature of natural selection believe that even if natural selection gives rise to a directional trend, it is a purely contingent occurrence in the sense of it being the result of a confluence of unrelated causal processes. We will argue against this view, and show how the challenge of causal complexity is relatively easily overcome, as long as we assume the organism-environment to be effectively stable (this assumption is common in population genetics).

When this assumption is dropped, it is much more doubtful whether natural selection can cause directional trends in evolution. If the same moth population is placed in a different environment, what was previously adaptive may no longer be so. In this way, this type of directionality also depends on contingent initial states, namely the state of the environment. If the environment changes, so does the direction of natural selection.[10] This is particularly relevant for trends in evolutionary history as a whole,

[10] As we will see in chapter 6, it is tempting to use the arms-race dynamic (or 'evolutionary escalation': Vermeij 1987) to argue why certain privileged adaptations (such as complexity and energy-intensiveness) are non-contingent outcomes of evolution. However, I will argue that such trends are still contingent on certain environmental states occurring more often than others.

since environments have undergone extreme changes over the course of evolutionary history.

3.2 Trends Arising from Phylogenetic Constraints

If the contingency of the first type of trend arises from changing external processes (i.e., changing environments), the contingency of the second arises from changing internal processes. This type of directionality, recognized by the Gouldian interpretation, arises from what Crick 1968 described as 'frozen accidents'. Originally Crick used the term 'frozen accident' to describe a hypothesis of how DNA did not originate for any special reason, but once it came into being it characterized all subsequent life. However, the term is used more generally to describe how a contingent event can nonetheless constrain subsequent evolution. Certain areas in possibility space are precluded and are made improbable, and in this sense the event introduces a 'directionality' into the subsequent evolution: certain paths become more probable than others.

We have mentioned such constraints on many places in this dissertation without going into much detail on *how* such constraints actually constrain evolutionary trajectories. The general phenomenon by which past evolutionary events can constrain subsequent evolutionary paths is termed *phylogenetic inertia*. In analogy to inertia in Newtonian physics[11], Hansen and Orzack describe it as "the tendency of a trait to resist a current adaptive force" (Hansen and Orzack 2005: 2063) – by reducing the available variation natural selection can act on.

However, phylogenetic inertia need not only refer to the evolution of a trait, but can refer to the way an evolutionary trajectory more generally can be constrained by previous adaptations, mutations or drift events (cf. Johnson et al. 1999: 759). Further, phylogenetic inertia can also refer to a tendency to resist drift or mutational changes. In this respect, McKitrick's definition 1993 of the related concept of phylogenetic constraint is the most general: "any result or component of the phylogenetic history of a lineage

[11] A pre-Synthesis (and refuted) sense of phylogenetic inertia is rectilinearity, or the tendency of phyla 'to evolve in one direction without deviation' (Simpson 1944b: 150). However, unlike in Newtonian dynamics, constant change is not equivalent to stasis in evolutionary dynamics. (See also McShea and Brandon (2010) for an argument that stasis, not change, indicates the presence of evolutionary forces.)

that prevents an anticipated course of evolution in that lineage".

In this way, phylogenetic inertia can be understood a very broad phenomenon but may refer to a variety of specific types of constraint. While such constraints are broadly compatible with the Modern Synthesis framework, as surveyed above, they tended to be ignored until the work of Gould and others in the 1980s. Today the biological constraint literature has become too complex to be briefly surveyed[12]; however, following Shanahan (2011), it is useful to distinguish between three basic types of constraint.

The first is genetic constraints: if some variations of the genotype are not present, then this will preclude certain evolutionary trajectories. For example, Antarctic fish do not have the genes that code for the proteins and regulatory mechanisms needed to live in warm waters; hence, the current genetic variation precludes an evolutionary path where some Antarctic fish species migrates to a warmer environment.

Such a constraint could be overcome by the introduction of a genotype into the population; however, this is not always possible due to underlying developmental constraints. A simple example of a developmental constraint is the difficulty for mutations in developmentally central genes to be viable – such as the Hox gene, which determines important features of the body plan (such as placement of limbs). The Hox gene is an example of a gene playing an important role in 'upstream' development, and when it would be changed, all the processes 'downstream' would need to be correspondingly changed as well. Hence the difficulty of modifying such upstream genes: the likelihood that a mutation would simultaneously change all those genes responsible for downstream developmental processes as well is extremely low (Shanahan 2011).

Such a constraint could be overcome by the introduction of a genotype into the population; however, this is not always possible due to underlying *developmental constraints*[13]. A simple example of a developmental constraint is the difficulty for mutations in developmentally central genes to be viable – such as the *Hox* gene, which determines important features of the body plan (such as placement of limbs). The *Hox* gene is an example of a gene playing an important role in 'upstream' development,

[12]See for example Antonovics and van Tienderen (1991) or Arnold (1992).

[13]In the framework of Arnold (1992), developmental constraints give a deeper explanation why genetic constraints exist.

and when it would be changed, all the processes 'downstream' would need to be correspondingly changed as well. Hence the difficulty of modifying such upstream genes: the likelihood that a mutation would simultaneously change all those genes responsible for downstream developmental processes as well is extremely low.

A final example of constraint – one that illustrates how difficult it may be in practice to distinguish between trends resulting from natural selection and trends resulting from 'constraints' – is fitness constraint. Fitness constraints occur when a certain phenotype cannot be realized given the current phenotype because the intermediary forms are maladaptive. Shanahan (2011) gives the example of kangaroo locomotion: even if bipedal running would be more adaptive for the kangaroo lineage, there would be no way it could evolve because the intermediary forms in between leaping and running are maladaptive. In this way, fitness constraints may be a consequence of underlying developmental constraints that prevent certain adaptive forms from appearing in a population.

4 The Challenge of Causal Complexity for Evolution by Natural Selection

In this way, the Modern Synthesis gives no theoretical reason to believe that any evolutionary path should be more probable than another if the tape of life were to be played from the beginning. Even while certain directions may arise through natural selection, these directional trends are likely to be short-lived, as they will be broken by changes in the environment. Even if they are not short-lived, they are contingent, and do not characterize all possible evolutionary histories. Directional trends may arise without natural selection, for instance through the action of evolutionary constraints, but very different constraints might have arisen had a slightly different evolutionary path been taken initially.

In this way, while the Gouldian and convergence-centric interpretations disagree as to how strong constraints are relative to natural selection, the fundamental concepts they both rely on imply that all directional trends are contingent. This is not so much a problem for the Gouldian narrative, as it is for the convergence-centric narrative, since it aims to show how certain outcomes are 'inevitable'.

This does not mean that there are no other resources in contemporary biological theory with which to overcome the challenge of causal complexity. In particular, it is increasingly being recognized that the phenomena of plasticity and niche construction, emphasized in the Extended Synthesis, significantly change our understanding of the causal structure of natural selection.

In the rest of part II, therefore, we will focus on natural selection in more detail in order to examine what types of patterns may be expected in subsequent evolution. In particular, the discussion will target two aspects of the challenge of causal complexity, each corresponding to two different aspects of the complexity of natural environments: the great number of independent causal processes that characterize organism-environment interaction, and the variability of these processes.

The first one casts doubt not only on the possibility of making a generalization about the types of trend that can be expected in evolution by natural selection, but also on whether natural selection can even be considered a causal process at all. When evolution by natural selection is analyzed at a sufficiently fine-grained level, what appears is that 'natural selection' is simply individuals producing varying numbers of offspring in a population. Natural selection then appears to be an arbitrary abstraction, projected onto individual-level causal processes for explanatory reasons. In no way does it cause population change, let alone directional trends. If directional trends occur, they are the result of a confluence of individual-level causal processes – in a way similar to how an unloaded fair die might still land on six multiple times in a row. This aspect of the challenge of causal complexity is discussed in the next chapter, where it is called the 'no privileged abstraction' argument.

The second challenge is that variability in the environment implies that natural selection will change its 'direction' along with the environment. This challenge has been alluded to in chapter 3 when discussing the way in which the outcome of convergent evolution is sensitive to initial environmental conditions. We will discuss this in more detail in chapter 6.

Chapter 5

Local Symmetries in Evolution by Natural Selection

In this chapter I will be examining the causal nature of natural selection, and what generalizations can be made about evolutionary patterns resulting from natural selection given the fact that many environmental processes may 'interrupt' natural selection. I will be situating this discussion in the context of the debate between the statisticalist and causalist interpretations of natural selection, and in the context of issues concerning how drift and natural selection can be distinguished.

Natural selection, both in biology textbooks and philosophical representations (e.g. Sober 1984), is often represented as some kind of Newtonian force, with magnitude and direction, originating in fitness differences and driving evolutionary change. By contrast, the metaphor has been rejected by the statisticalist view of natural selection (e.g. Matthen and Ariew 2002, 2009; Walsh et al. 2002), which claims natural selection is a mere book-keeping device, a way to keep count of the genuinely causal interactions that take place between individual organisms. Not even a cause, selection is an epiphenomenon, useful only for explanatory purposes.

While there are multiple ways of parsing this specific debate, the approach I will be adopting – consistent with that in part II – is to examine how the statisticalist view arises from the *complexity* of the organism-environment interaction and from the ensuing *reference class* problems. As mentioned in the previous chapter, complexity here refers to the very

large number of independent processes that affect how often an organism will have reproduced by the end of its life, and the reference class problems that arise from this reflect the absence of any privileged abstraction away from this complexity.

Such themes have been noted before (for example, Brandon 1990, 2005; Kaplan 2013; Strevens 2016), but remain, on the whole, underemphasized. On the basis of a few assumptions about organism-environment interaction, a principled argument for the statisticalist view can be reconstructed, somewhat in the following vein (to be laid out in detail later): most of the individual-level interactions that affect reproductive outcome are abstracted away from in an explanation by natural selection, but it is the complex totality of organism-environment interactions that in fact causes populations to change. Hence natural selection is an epiphenomenon without any causal impact, a statistical apportioning of the fundamental causal reality.

Some of the main causalist counterarguments fail to make inroads on this ('no-privileged-abstraction') argument. One such counterargument is that natural selection is a cause insofar as intervening in natural selection at the population level 'makes a difference' to how populations evolve (Reisman and Forber 2005; Millstein 2006; Shapiro and Sober 2007; Glennan 2009; Gildenhuys 2014). Examples of interventions include modifying the selection coefficients and fitness values (Forber and Reisman 2007; Shapiro and Sober 2007), or introducing new phenotypes (Millstein 2006). This counterargument was rejected by Matthen and Ariew (2009) and Walsh (2007; 2010), who pointed out in various ways that the correlations between natural selection and different population parameters are mathematical in nature, not causal.

Furthermore, attributing fitness values to a group of individual organisms already presupposes a certain abstraction away from the complexity of organism-environment interaction, and thus such counterarguments beg the question. Hence, the question to be focused on instead is: are the abstractions involved in explanations by natural selection problematic for selection's supposedly causal nature?

I will seek to undermine the no-privileged-abstraction argument by considering a class of counterexample: tendencies towards stable equilibrium. A population tending towards stable equilibrium is strong evidence

for natural selection (Endler 1986); such a tendency also allows certain abstractions to be privileged over others. In particular, the individual-level processes, which do not make any causal difference at the longest timescales (over multiple generations) and thus do not make a difference to the outcome state,– may be abstracted away from. In such cases natural selection may be understood as a force, with definite direction and magnitude, driving populations towards equilibrium.

Not all instances of evolution by natural selection actually tend towards stable equilibrium; in particular, populations undergoing frequency-dependent selection may fail to settle down at all. Such cases provide a challenge to the equilibrium model of selection defended in this chapter, and towards the end of the chapter I will outline two possible ways in which they can be dealt with.

Nonetheless, perhaps the most general consequence of the account presented here is that it implies an alternative (or modified[1]) metaphor to that of Newtonian force. Natural selection is closer in nature to entropic forces, which are non-fundamental forces that originate in a system's statistical tendency to evolve towards a configuration with higher entropy. While the concept of entropy may not be applicable to evolution by natural selection (populations are far-from-equilibrium systems), natural selection, like entropic forces, and in contrast to typical Newtonian forces, does not have a localizable spatiotemporal source, and does not produce movement but a reconfiguration of the system. Some on the causalist side of the debate may be uncomfortable with the suggestion; however, entropic forces are not epiphenomenal: they have real consequences (e.g. osmosis, the Casimir effect), and the distinction between fundamental and entropic forces may not be so clear-cut at the level of fundamental physics (e.g. Verlinde 2011).

The chapter is structured as follows: in the first section I will make a principled case for the statisticalist position, and in the second section will criticize it in light of how field biologists test the presence of natural selection. In the third and fourth sections I present a model of natural selection as a causal tendency towards stable equilibrium, and deal with potential objections. The final section is a discussion about force metaphors, and

[1]There is no consensus as to what precisely are the defining conditions (e.g. additivity) of Newtonian forces. A detailed discussion of this would bring us too far from the purposes of this chapter the topic at hand, but see Wilson (2007) or Stephens (2010).

about the limitations of the account.

1 The No Privileged Abstraction Argument

The interaction between an organism and its environment is, generally speaking, complex. More often than not, environments are chaotic, characterized by ever-changing weather, geography and ecology. An organism itself is composed of many interacting parts, across numerous hierarchical structures – organ systems, organs, tissues, cells, nutrient gradients, etc. The resulting interaction between organism and environment has such a vast number of degrees of freedom as to make any parametrization practically impossible. Interactions can be biotic, such as competitive, symbiotic and parasitic interactions, or abiotic, such as extracting resources from the environment, or being affected by fundamental forces in the environment. While ecological models focus on only a few of these interactions at a time, in general, we can assume that the number of degrees of freedom characterizing the actual organism-environment interaction is very large.[2]

Yet even among these difference-making interactions between organism and environment, many will be abstracted away from in explanations by natural selection. The vicinity of a moth to a forest fire is likely to be ignored — unless, of course, the moth possesses some (perverse) heritable trait that makes it more likely to seek out fires. Once-off interactions are ignored, and interactions linked to the traits of an organism are included. On what basis are such traits selected, and what is their relation to the fitness (understood as the expected reproductive outcome[3]) of the organism?

The relation between trait and organism fitness is a difficult one — and, judging from a recent exchange between Sober (2013) and Pence and Ramsey (2015), the two concepts seem intertwined. Pence and Ramsey have argued that definitions of trait fitness depend on organismic fitness, roughly because relevance for organismic reproductive success is the only

[2]To emphasize: complexity is here taken to refer to the number of degrees of freedom, not functional complexity, or number of part-types (see Strevens 2003; McShea 2000).

[3]There are more accurate and sophisticated measures of fitness available (see Pence and Ramsey 2013), but how precisely fitness values are extracted from life histories does not materially affect the argument of this chapter.

measure by which to judge the relevance of a trait. By contrast, Sober has argued that organismic fitness is an actuarial quantity, estimated on the basis of how various traits affect organismic reproductive outcome.

Regardless of which fitness concept is the more fundamental one, statisticalists have denied that either measure of fitness can be estimated in a non-arbitrary way. Matthen (2009) has proposed that fitness measures are obtained by means of the condition of *metaconstancy*: a process may be excluded from consideration only if it is neutral with respect to (i.e. probabilistically uninfluenced by) all *heritable* properties. There is no heritable trait of a moth that correlates with its proximity to forest fires, and hence the evolutionary biologist may ignore it. Thus, in explanations by natural selection the causal difference-making processes that do not correlate with a heritable trait are to be ignored (see Figure 5.1).

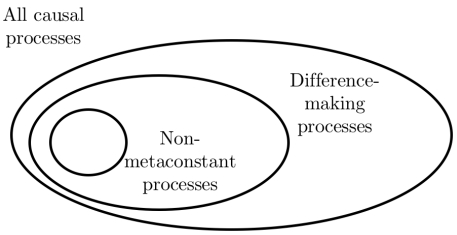

Figure 5.1: Fitness abstracts away from most causal processes.

The consequence for individual fitness is that there is no privileged way of calculating the probability that an organism X will have n offspring. If we exclude all non-metaconstant processes, we may arrive at a probability p; if we take some other non-metaconstant process (*e.g.* the strength of the wind) into consideration, we might estimate a probability $p' \neq p$. The value p reflects a particular theoretical interest, not an objective propensity. Yet it is the probability distribution (over possible reproductive outcomes) that takes into account *all* difference-making processes that corresponds to how the population *actually* changes from generation to generation. Individual fitness, as it is usually defined, only picks out one particular aspect of population change.

Trait fitness, taken here as the expected reproductive outcome of an organism given a trait, does not fare much better[4]. The probability of an individual having n offspring given trait T may depend on an indefi-

[4]One can also define trait fitness as the expected number of duplicates in the offspring generation. However this definition runs into similar problems due to its dependence on organismic fitness (see also Pence and Ramsey 2015).

nite number of other *unspecified* traits $\mathrm{T}_1, \mathrm{T}_2, \ldots$ and so on — individual organisms are by presupposition complex entities. According to which of the other traits $\mathrm{T}_1, \mathrm{T}_2, \ldots$ an organism (of the same species[5]) possesses, a different probability will be obtained. There is no privileged way of calculating trait fitness either.

In a way this argument is fundamentally a restatement of the reference class problem in a biological context. Individual organisms can be classified into groups in many different ways, due to the complexity of the individual-environment interactions, and hence there is a real problem in estimating the probabilities underlying the actual births and deaths.

One may think a simple way out of the reference class problem for trait fitness would be to exhaustively specify the traits of X, and to estimate the causal impact of T by comparing X to an identical organism that is missing T. Thus the reference class of X comprises the individuals with T in a hypothetical, infinite population where all the other traits $\mathrm{T}_1, \mathrm{T}_2, \ldots$ are equally represented. Cannot one thus arrive at a privileged calculation of $P(\mathrm{w} = n|\mathrm{T})$? Selection 'for' camouflage would then be represented by the quantity $s = E(\mathrm{w}|\mathrm{T}) - E(\mathrm{w}|\neg\mathrm{T})$, where $E(\mathrm{w})$ represents the expected reproductive outcome. In principle this would indeed be possible; however, the problem is then that the quantity s may not describe natural selection in the *actual* population, where many of the traits $\mathrm{T}_1, \mathrm{T}_2, \ldots$ may be absent.

The no-privileged-abstraction argument allows some key statisticalist positions to be reconstructed:

(1) Fitness can only be estimated through regression. Since some difference-making processes are excluded from the explanation by abstraction, there is no a priori way of estimating their impact on the probabilities of surviving and reproducing. Is the probability of having two offspring given the trait of camouflage .9 or .5? The value of that probability will be affected by the various causal processes impacting the individual, and these impacts will necessarily remain unknown since they were deliberately abstracted away from (see Matthen 2009). Fitness values can be estimated only retrospectively, by a statistical regression on the actually occurred births and deaths in a population. Only by extrapolating from past indi-

[5]We will not consider the question whether organisms of a given species can even be associated with a set of traits as this is outside the parameters of this investigation.

vidual events does fitness have some predictive value. In this way, fitness values are mere statistical descriptors, not causes of actual frequencies of births and deaths.

(2) Fitness components are not additive. The probabilistic relevance of having both traits T_1 and T_2 cannot be predicted from the relevance that T_1 and T_2 have by themselves, because the traits may be causally dependent (Matthen and Ariew 2002; Matthen 2009). For example (adapted from Stephens (2004)), having resistance to malaria may increase fitness by p and being able to run fast may increase fitness by q, but in general we cannot know what fitness an organism will have when it has both traits, due to the unknown interaction between the two traits. Selection is not a causal force that can be simply divided into components.

(3) Natural selection refers to the mathematical relation between change in frequencies and variance in rates, exemplified by Li's theorem. Selection is variation in fitness, and since fitness is only a statistical descriptor of actual population, selection is not distinct from population change. To say a population is changing and to say individuals are reproducing at differential rates is to say the same thing. The upshot is that selection is not a cause mediating between fitness differences and population change in the way that the gravitational force mediates between mass distribution and acceleration of bodies (as represented in Sober 1984); rather it is more like the shadow of a flagpole (Walsh 2000; Matthen and Ariew 2009).

This is also why causalist appeals to interventionist causality fail (Woodward 2003; Reisman and Forber 2005; Millstein 2006; Shapiro and Sober 2007; Glennan 2009): while it is conceivable to intervene on the gravitational force alone while keeping mass distribution constant (for example by changing the value of the gravitational constant), it is not conceivable to manipulate selection without changing fitness differences. Neither does the argument imply that there are no population-level causes (as claimed by Shapiro and Sober (2007)): the epiphenomenality of selection is not a consequence of its population-level nature, but of its mathematical nature.

In summary, the no-privileged-abstraction argument can be represented in the following format:

(1) The space of life histories Ω is characterized a high number of degrees of freedom, where each independent variable makes a difference for the reproductive outcome of a life history.
(2) Fitness (expected reproductive outcome) is specified by ignoring all degrees of freedom except a select few, dependent on explanatory interests.
(3) Yet evolutionary change, being constituted by individual births and deaths, is affected by *all* degrees of freedom of Ω, and is not a subset of them (from (1)).
$\therefore$ (4) Natural selection, defined as variation in fitness, describes a particular pattern in population change, but since it does not refer to the processes that define Ω, it does not correspond to the cause of population change (from (2) and (3))

2 Epistemic Routes to Natural Selection

In light of the preceding argument, the central problem becomes an epistemological one: can selection (and drift) be estimated from complex, natural populations – and not idealized toy models – in such a way that we can be *justified* in believing that selection is an objective cause, and not just one way among many of statistically apportioning the underlying causal reality?

While a general solution to the reference class problem is likely impossible (Hájek 2007), I will argue there is a way out in the particular case of evolution by natural selection. In this section I will lay the ground by arguing that previous causalist arguments based on manipulating abstract models do not work, and that instead we should examine how field biologists test for selection in natural populations.

2.1 Abstract Models

One way the causality for selection has been argued for is by taking deviations from certain null models, such as Hardy-Weinberg equilibria, as evidence of the causal impact of selection (e.g. Sober 1984; Stephens 2004). Further, since deviation from those models indicates the presence of selection, it is sometimes claimed that natural selection must be an objective

cause of evolutionary change, and not merely some arbitrary statistical redescription of individual births and deaths (Gildenhuys 2014).

However, I would argue against the claim that "departure from Hardy-Weinberg equilibrium is sufficient (rather than necessary) for thinking that some force must have been at work" (Stephens 2004: 559), for two reasons. First, to claim this is to ignore the fact that Hardy-Weinberg equilibria heavily abstract away from the complexity of individual-environment interactions, and that they do not necessarily capture change in real populations, even if there is no natural selection at work (see also Brandon (2006)). Some of the presuppositions of the Hardy-Weinberg null model are that individuals reproduce sexually, mate randomly and do not overlap with their offspring. Even when the test is properly applicable, Endler (1986: 65) notes that it is a weak test statistically and that by itself the method cannot demonstrate selection.

Second, on a more fundamental level, even if the presence of natural selection were to be established despite the difficulties, it still would not be able to adjudicate between a statisticalist and a causalist interpretation of natural selection, and would not establish the presence of a 'force'. The ambiguity in divvying up the individual-level causal processes would remain. For example, if the frequency of the recessive homozygous genotype aa is higher than expected by the Hardy-Weinberg model, this only means that the genotype aa is correlated with a larger number of offspring in the population under consideration. It may turn out that some other allele B is not present at all in the population, and if it were, that different correlations between aa and offspring number would be obtained (see discussion about trait fitness in previous section). Thus the test does not establish an objective fitness differential between aa and the other traits.

Another causalist strategy has been to draw on direct estimates of selection, through response to selection (by means of the breeder's equation) or selection coefficients. Manipulating selection coefficients correlates with population change, and hence is taken as a sign that selection must be causal (e.g. Forber and Reisman 2007). However, even here the ambiguities cannot be avoided, since they turn up in the manipulation of the coefficients. Coefficients are not manipulated directly: what is manipulated in practice are traits, trait frequencies and trait-environment interactions, and that such manipulations lead to population change is perfectly compat-

ible with an epiphenomenalist interpretation of the selection coefficients. Whether or not changes in selection coefficients can be neutrally estimated on the basis of such manipulations is precisely the point of contention.

Even the arguments used by Sober in his well-known distinction between selection-of and selection-for rely on prior abstraction. His well-known illustration of the distinction between selection-of and selection-for relies on a toy model, where organisms (marbles) have two properties (colour and size) and have a single difference-making interaction with their environment (size of the hole). Thus selection for size seems to cause the change in distribution in the marble population. The problem, however, is that real-life organisms possess many more than two traits, interact with the environment in more than a single way (fitting through a hole), and that both the number of relevant traits and relevant interactions is unknown. The question at hand is whether such abstract models can be given a causal interpretation; hence they cannot be relied upon to defend a causal interpretation of natural selection.

2.2 Natural Populations

While abstract models cannot be used to craft a response to the no-privileged- abstraction argument, it is helpful to consider the methods used by field biologists to establish the presence of selection. Such methods cast doubt on whether truly there is not any non-arbitrary way out of the reference class problem. Of course, whether or not these methods are sufficient to establish selection as a *cause* of evolution is a separate question, and will be considered in the next section.

In actual field studies, estimating selection coefficients is often very difficult because of the variability in the temporal dynamics of population change. Siepielski et al. (2009), in a review of replicated field studies, indicate that, in general, natural selection varies in strength, direction (positive and negative) and form (linear and nonlinear selection) from one generation to the next.

This variability in selection dynamics is a reflection of the complexity of individual-environment interaction. Environments are rarely static, and slight changes in biotic or abiotic variables may have large effects on the patterns of births and deaths exhibited. Ideally we should be able to exactly replicate population and environmental structure as to establish

whether population changes are simply noise or follow an expected pattern; however, in the vast majority of cases this is not possible. Even when adequate temporal replicates exist, the duration of the study is very rarely longer than a couple of years (Siepielski et al. 2009).

In such cases, where variability is so ubiquitous, it becomes difficult to rebut statisticalist challenges. Many of the assumptions in population genetics models (infinite populations, static environments, absence of linkage, and so on) are made in order to gain mathematical tractability. However, while variability and complexity may be eliminated in abstract models, they are more or less permanent features of real environments, and one may legitimately question whether such abstract models can be used to judge on questions of causality when it is unclear whether even the basic inputs of the models (such as selection coefficients) reflect objective properties of organisms and populations.

Nonetheless, there is a powerful group of methods for cutting through short term and individual-level complexity: tests for stable equilibrium. Endler (1986) describes two ways in which equilibria may be used to establish the presence of natural selection. The first method (method V) requires longitudinal studies on trait frequency distributions, and tests either for long-term stability in frequencies, or steady directional change. Of the two, an observed directionality is taken as stronger evidence for natural selection. However, it is not foolproof: Endler describes how simulations of coin-toss runs surprisingly often give rise to (temporary) directional trends. Even though a single unbiased coin-tossing experiment leads to heads as often as it does to tails, the probability of a significant run of either heads or tails is surprisingly high.

The second method (VI) has a manipulative character: a population at stasis is perturbed, and a subsequent observation of directional change, either back to its previous state or towards a new state, is taken as strong evidence for natural selection. One implementation of the perturbation entails manipulating the trait frequencies of the population. Reversion to the previous distribution of frequencies is taken as strong evidence for the presence of natural selection. Changes in the environment constitute another type of perturbation. Such perturbations can be human induced (e.g. pesticides) or natural (e.g. volcanic eruptions, epidemics), and can occur abruptly or gradually. An example of where a gradual change in environ-

ment is used to establish selection is the well-known study of moth evolution in response to increasing air pollution in Britain (Kettlewell 1955).

Such methods are not always possible. Sometimes there are multiple equilibria in the system, complicating the interpretation. Further, not all forms of natural selection, such as certain types of frequency dependent selection, seem to tend towards equilibrium. Nonetheless, when there is a tendency towards equilibrium, it can be discerned in longitudinal studies as a long-term effective change in population structure, despite short-term fluctuations.

3 The Equilibrium Model of Causality

While such methods undercut the generality of the no-privileged-abstraction argument, in themselves they do not constitute a direct argument for the causality of selection. In this section such a direct argument will be attempted. We will assume that a natural population is undergoing *directional change* in its trait frequency distribution, and is evolving towards a stable equilibrium. This assumption may be thought of as the outcome of an empirical field investigation. In the next section the assumption will be discussed more critically, especially with regards to the worry that this introduces some arbitrary abstraction.

I will first argue that natural selection, as tendency towards equilibrium, is a cause according to the (probabilistic) counterfactual understanding of causality (e.g. Lewis 1986); subsequently I will argue it is a force with direction and magnitude.

3.1 Relation and Relata

Let us first briefly outline the landscape of accounts of the causality of selection with regard to two reference points: the causal relation and causal relata. This will allow for the view defended here to be situated with more precision.

Concerning relation, I subscribe to the statisticalist criticism that the relation is not one of causal *production* (Matthen and Ariew 2002): evolution by natural selection is not a spatiotemporally continuous process where some physical mark is transmitted, or quantity conserved (Salmon

1984; Dowe 2000). However, most causalist accounts have drawn on a difference-making relation. In light of problems arising from the use of the interventionalist criterion of causality to adjudicate on the epiphenomenality of population-level processes (Matthen and Ariew 2009; Baumgartner 2010), and in light of the fact that such interventions often depend on abstract models, I will avoid that specific conception of causality (as does e.g. Huneman 2012). Instead, I will consider whether natural selection represents a probabilistic *counterfactual* relation between its relata (e.g. Lewis 1986).

Even within the counterfactual option there are a number of possible accounts according to how the relata — the source and effect of natural selection — are to be understood. One point of controversy concerns the *level of analysis* at which the source and effect of natural selection are to be located. Bouchard and Rosenberg (2004) propose that the source of natural selection consists in pair-wise individual-level competitive differences and that the effect consists in pair-wise differences in (expected) individual reproductive outcomes (Bouchard and Rosenberg 2004). I follow Millstein's (2006) criticism of this proposal, and consider natural selection at the population level.

A further point of contention is whether the source of selection is (population-level) fitness differences and whether the effect is changes in trait frequencies (e.g. Sober 1984; Millstein 2006). One problem here is that natural selection and changes in trait frequencies are mathematical, not causal, consequences of fitness differences (Matthen 2009). Another problem, as argued previously in this chapter, is that it is not clear whether fitness variables are objective in the same way as mass or acceleration, and hence to claim selection is causal because fitness values can be manipulated is to beg the question.

By contrast, the proposal of this chapter can be summarized in the following scheme:

> population with trait distribution $\mathbf{d}$ and in environment $E \rightarrow$
> natural selection $\rightarrow$ directional change in trait frequencies

The vector $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ specifies the relative frequency of each of the n traits that characterize organisms in the population. I will now argue

that this proposal ensures that each of the relations (between selection and source, and between selection and effect) is causally counterfactual.

First, with regard to the former relation, the combination of n traits and environment E gives rise to a great number of biotic and abiotic organism-environment interactions. Some of those interactions will make no net difference to the effective direction of population change, while others will. Natural selection is only constituted by the latter interactions.[6]

Understood in this way, natural selection is not a mathematical consequence of $\mathbf{d}$ and E because the causal interactions between organisms of the population and between organisms and environment are not mathematical consequences of a specification of $\mathbf{d}$ and E. (For example, if the laws of physics were changed, the same $\mathbf{d}$ and E would give rise to different causal interactions.) Further, if $\mathbf{d}$ and E were different, we would likely see different causal interactions, and thus a different process of natural selection.

Second, the effect of natural selection is to be sought in the *direction* of change, not the fact of change. This proposal contrasts with the accounts presented by Millstein (2006), Matthen and Ariew (2009), and Huneman (2012), who all locate the effect of natural selection in the *fact* of population change. The problem with the latter proposal is that population change is just as likely to occur where natural selection is absent — such as when drift is present, or when the population evolves towards a Hardy-Weinberg equilibrium state. Hence the question to be asked is not whether selection causes an accumulation of births and deaths, but whether selection causes a stable trend in the accumulation of births and deaths.

One could compare the effect of natural selection on a population to the effect of an electric field on the random walk of a charged particle in a gas. The gas has a certain temperature, and the particles collide randomly; however, because of the electric field there is a bias in a certain direction. The field does not cause the movement of the particle itself, but does make a difference for the long-term 'effective' path the particle will take (Figure

[6]This is continuous with Huneman's (2012) definition of selection pressures as ecologically "reliable factors which differentially affect the trait types" (185). The only difference is that reliableness is specifically defined here in reference to an effective directionality.

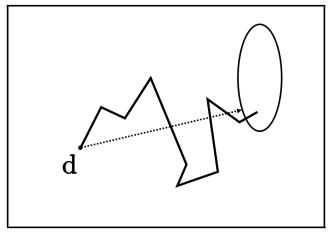

Figure 5.2: Actual versus effective evolutionary change.

5.2). In this way natural selection's causal influence is to be sought in the effective direction of population change instead of change itself, as the latter could also be the result of random fluctuation (drift).

The relation between natural selection and effect is causally counterfactual: were natural selection not present, directional evolution would be much less likely. Drift could in principle give rise to directional evolution, but the probability of this occurring diminishes the longer the directional change lasts. Further, natural selection is not sufficient for a directional change in frequencies to take hold, as natural selection can be counteracted by drift.

At this point, one may rehearse an objection, going back to Sober (1983), that equilibrium explanations are not causal (see also Huneman (2010)for a generalization of this argument). Sober argued that explanations where the outcome state is explained as a stable equilibrium state are not causal because such explanations do not pick out the *actual cause* of the outcome state. For example, if a marble is let go at the rim of a bowl, and proceeds to find its way to the lowest point of the bowl, we explain its outcome state by the shape of the bowl, not by the particular path the marble followed. In other words, we explain by referring to a disjunction of possible causal scenarios (1983: 84): if a system had not followed this particular pathway to equilibrium, it would have attained equilibrium by another path.

However, this does not present a problem, because Sober is drawing on a different notion of causality. If we mean by cause the actual, specific path a population takes on the way to equilibrium, then indeed an explanation by natural selection does not pick out any actual cause. However, a counterfactual difference-making relation exists between the equilibrium state

and the directionality of the *trend* leading up to the equilibrium. Explanations by natural selection are causal insofar as they pick out directional trends as causes of the outcome state (equilibrium).

A second objection concerns the directional evolution induced by the two competing models of zero-force laws: Hardy-Weinberg (H-W) equilibria (Sober 1984) and drift (Brandon 2006). Do the population dynamics described by drift or by the H-W equilibria not lead to equilibrium states?

If only drift acts on a population, then indeed the population will continue to change its structure until one trait becomes fixed. However, the absorbing state of fixation is not a stable equilibrium, but a neutral equilibrium state. When perturbed from the state of fixation of some trait, the population will not necessarily drift back to the same state, but may drift to the fixation of some other trait. The process of drift towards fixation is like a drunk man's walk on a sidewalk with a gutter on either side. Once the drunk man falls into the gutter he stays there, but does not necessarily return to the gutter if pulled out by someone. There is no tendency towards falling into one particular gutter.

In an analogous way, an H-W equilibrium state is not stable either. If (p, q, r) is a H-W equilibrium state (p and r represent the relative frequencies of the homozygous traits AA and BB; q represents the relative frequency of the heterozygous trait AB), and is perturbed to $(p + \delta_p, q + \delta_q, r + \delta_r)$, then this perturbed state will only evolve back towards (p, q, r) if $\delta_p = \delta_r$ and $\delta_q = -2\delta_p$ (see appendix). H-W equilibrium points are not stable equilibrium points, as they are stable only along one specific line in state space.

In this way, if natural selection were not present, regardless of which zero-force model one adheres to, there would not be a trend towards stable equilibrium. Drift may lead to population change, but only selection leads to a robust, multigenerational trend in the change of population structure.

3.2 Direction and Magnitude

In instances where natural selection causes a population to trend towards a stable equilibrium state, the stable equilibrium is a reference point that allows a direction to be ascribed to the population change. Note that, stabilizing, disrupting and directional selection all are 'directional' in this sense, since they tend either towards the fixation of a particular trait, or

towards stable polymorphism.

The magnitude of selection can be estimated in the regular way, through the response to selection. The only difference in the equilibrium framework is that what matters is the *effective* response to selection. In the case only a single trait is selected, the breeder's equation is applicable.[7]

$$R_{eff} = H^2 s_{eff}$$

where H^2 is the heritability, R_{eff} the effective response to selection, and s_{eff} the (effective) strength of natural selection.

Thus, if a population never reaches the equilibrium state ($N \to \infty$), the effective response to selection is zero, even though the magnitude of selection may be nonzero. In this case the tendency towards the equilibrium state does not translate into any actual trend. Some underlying reasons for this can be seen by separating heritability into genetic and environmental variation: $H^2 = V_G/(V_G + V_E)$ where V_G is the portion of total phenotypic variance that can be explained by genetic variance, and V_E is the portion explainable by variation in the environment. Thus, as drift increases, the environment will have an increasingly variable effect on the population, and hence the effective response to selection goes to zero. Likewise, if traits are transmitted poorly between generations, the response to selection will be diminished.

4 Statisticalist Objections

In this section we consider two more fundamental objections to the model. The first is that it simply reintroduces a different arbitrary abstraction away from the complexity of organism-environment interaction. The second is that natural selection still has an epiphenomenal character in the model.

[7] In the case of multiple traits, this is generalized to Lande's equation: $R_{eff} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s_{eff}}$, where $\mathbf{G}$ and $\mathbf{P}$ are the additive genetic and phenotypic variance-covariance matrices.

4.1 Directionality and Abstraction

Even limiting attention to cases where populations trend to stable equilibrium, one objection could be that defining natural selection with respect to the effective direction of population change is simply a different way of arbitrarily abstracting away from the complexity of individual-level causal processes. Based on an observed approach to stable equilibrium, the equilibrium model distinguishes between processes that make a difference to an individual's reproductive outcome (Δ_2) and those that make a difference to the stable equilibrium (Δ_3 — see Figure 5.3). However, or so the objection goes, in principle some explanatory interest in some other feature of the population change could lead to some different distinction being made among causal processes (Δ_2' and Δ_3'), and some other direction (if any at all) being ascribed to the population change.

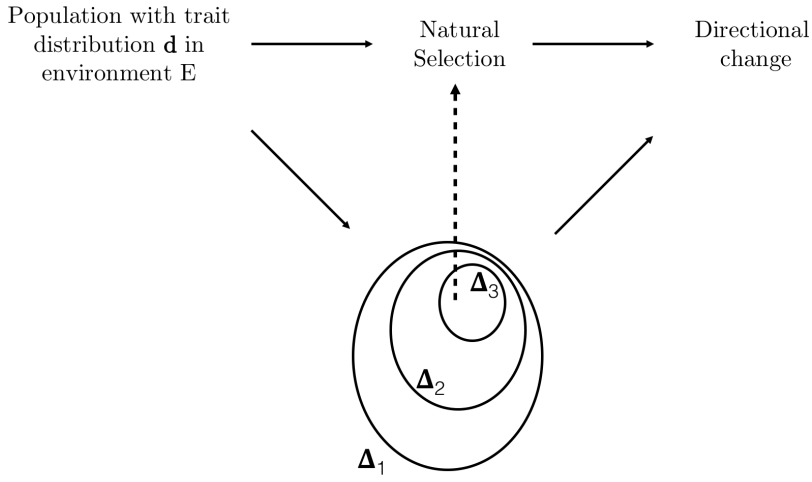

Figure 5.3: The solid lines represent causal relations; the dashed line represents a constitutive relation. Δ_1 represents the set of all individual-level causal processes (defining a specific life-history), Δ_2 the set of all individual-level processes affecting the reproductive outcome of a life-history, and Δ_3 the processes which make a difference to the equilibrium distribution of trait frequencies.

In response, one must first note that some dependence on explanatory

interests is unavoidable. When observing the behaviour of natural populations, an evolutionary biologist will be interested in explaining different types of outcome than, say, a biochemist, or an insurance agent. However, what is at stake is the implication of the no-privileged-abstraction argument, that even when the outcome of interest is decided on, there is no privileged way of (causally) explaining the evolutionary change in virtue of the complexity and many-angled nature of organism-environment interaction.

The equilibrium model challenges this latter claim, arguing that sometimes population-level behaviour is *simple* even though the organism-environment interaction may be *complex*. The reason why simplicity can emerge out of complexity is simply because many individual-level interactions 'average-out' to zero, which is to say, they make no long-term difference to the structure of the population (see also Strevens 2003). At a given time-scale, there is an objective fact of the matter, independent of the type of trait or interaction one happens to be interested in, as to which individual-level interactions make an overall difference and which do not.

Beyond choice of outcome state, the only dependence on explanatory interests allowed for by the equilibrium model is that on *time-scale*. What makes a difference in the long-term does not necessarily make a relevant difference in the short-term, and vice versa. Further, there might not be any processes which make a difference in the long term: in such cases, complexity of organism-environment interaction does not give rise to a simple approach to equilibrium (see next section). However, there is no dependence on the traits or individual-level interactions the observer happens to find interesting.

4.2 Epiphenomenalism

One of the key statisticalist challenges was to point to the epiphenomenal character of natural selection: selection is a '*tertium quid*' that could be eliminated from consideration without affecting the causes actually driving evolution. The original version of this challenge focused on the mathematical character of natural selection, and while we have already argued why natural selection is not a mathematical consequence of the source of selection, one could still object that the causal relations connecting organism-environment structure, natural selection and directional change could be

eliminated without affecting directional change.

Figure 5.3 schematically represents this objection: if natural selection were to be eliminated from consideration, the causal relation between the processes determining individual births and deaths, Δ_2, and directional change would still be intact. Evolution by natural selection is still 'just' an accumulation of births and deaths in the equilibrium model.

In response, one need only to point out that only the processes in Δ_3, not Δ_2, cause *directional* change. Aspects of the causal processes impinging on organisms in a population may be changed without changing the long-term outcome, but the latter depends counterfactually on the long-term difference-making processes (Δ_3). To eliminate natural selection from the causal scheme is to eliminate the difference between Δ_2 and Δ_3—and this affects the causal scheme[8].

A related way of putting the same point is that natural selection contains *more* information about the causal structure of population change than is contained by the mass of causal processes Δ_2. Unlike Δ_2, natural selection tells us something about *possible* evolutionary changes, not just about actual changes. Specifically, it tells that if the population were to be perturbed, the population would tend to return to equilibrium, in virtue of the processes in Δ_3. Eliminating natural selection would eliminate this information about the causal structure of population change.

5 Discussion: General Implications

5.1 What about Frequency-dependence?

Throughout the chapter we have alluded to cases such as frequency-dependent selection, where evolution by natural selection does not always tend towards stable equilibrium. Strictly speaking such cases are not covered by the account presented here, and in principle the equilibrium model could be interpreted as having limited scope, unchallenged by cases such as frequency-dependence. However, it is natural to wonder to what extent the model could be extended to cover such cases also.

[8]In this respect, the account presented here is continuous with the suggestion that natural selection is a structuring cause in the space of population life-histories (Ramsey 2015). See also Kaplan (2013) for the suggestion that natural selection is some measure of robustness.

While a proper treatment is beyond the scope of this chapter, an easy extension of the model would be to distinguish between a variable component (changing direction) and an effective component (fixed direction) of natural selection. The effective component would describe approaches to stable equilibria, and in general, attractor states, while the variable component would describe the rest. This distinction allows the question to be rephrased with some more precision: to what extent is the variable component of selection vulnerable to the no-privileged-abstraction argument?

Two alternative responses can be distinguished. The first would be to admit that sometimes it is less meaningful to apply the predicate 'causal'. When evolution by natural selection tends towards stable equilibrium, the process is clearly causal; however, as the overall directionality becomes increasingly difficult to recognize, to the point that population change continually changes apparent direction from one generation to the next, selection becomes increasingly indistinguishable from drift, and it becomes increasingly meaningless to attribute causality to natural selection. Thus instances of selection are to be situated on a spectrum, from clearly causal at one end (tendency towards stable equilibrium) to indistinguishable from drift at the other.

One could formulate this in terms of time-scale relativity. Viewed from a long time-scale, a population that cycles indefinitely might as well not be changing at all: the frequency dependent selection responsible for the cycling makes no causal difference at a long time scale. At a shorter time-scale, perhaps over tens of generations, selection does make a difference, and could be considered causal at this time scale, even though it will be more difficult to disentangle selection's causal contribution from drift. As the time-scale decreases to a single generation, it becomes increasingly difficult to judge what the precise causal impact of selection is (e.g. is an observed and non-replicable change in the frequency of a trait due to selection-for, selection-of, or drift?). Thus it could be ventured that it makes increasingly less sense to attribute causality to selection.

The other response is to reduce the case of variable selection to the equilibrium model by analyzing a single process of frequency-dependent selection as a succession of different selection processes, each tending towards (but not realizing) a stable equilibrium. For example, consider frequency-dependent selection for colour in populations of *Poecliid* fish (cf. Huneman

2012): predators are tuned to prey on fish with majority colours, so that any one colour will undergo cycles of selection-for and selection-against, depending on whether it is a majority or a minority colour. Is it the same instance of natural selection doing the selecting for and the selecting against? According to this response, it is not, as the interaction between a fish with a given colour and its environment changes in the two cases. Any cycle of selection-for and selection-against is a succession of selection processes each tending towards a different stable equilibrium.

To fully elaborate this response, one would need to discuss the problem of identifying the selective environment of an organism. Instead, I refer to an existing discussion by Brandon (2005), who argues that fitness can only be ascribed to an organism or trait if the selective environment is 'homogeneous' within the relevant region. If the selective environment is heterogeneous, then the nature of the organism-environment interaction fluctuates significantly, and Brandon terms the process of selection that arises from such an environment 'compound selection'. Similarly, frequency-dependent selection could be understood as compound selection, where the selective environment is temporally heterogeneous (due to changing biotic interactions).

In this way variable selection is reduced to a succession of different instantaneous selection processes, each with a stable direction. So, while the overall process of frequency-dependent selection may not be ascribed a direction, the *instantaneous* selection processes can – and hence these may be considered causal. Each instantaneous selection process tends towards equilibrium, even though equilibrium is never reached because the environment is continually changing.

5.2 What Kind of Force is Selection?

Finally, in order both to broaden the perspectives of the chapter and to bring it to an intuitive close, I would wish to consider natural selection on a more metaphorical level. I have argued natural selection is a causal force on the basis of certain formal characteristics (counterfactual dependence, direction, magnitude); however, it is fair to ask what kind of a force it is precisely. It is clearly not a classical Newtonian force: neither its source, nor its effect is spatiotemporally localizable. There is also no 'field' of force associated with natural selection. Is it possible to say something

more positive about what kind of force it is?

In their 2002 paper, Matthen and Ariew compare evolution by natural selection to heat flow (Matthen and Ariew 2002). During heat flow there is no transmission of a physical mark; it is instead a statistical process where the concentration of fast-moving molecules in the warm part of the gas spreads throughout the container. While the net effect is that energy diffuses smoothly, in reality this occurs one discontinuous collision at a time. Similarly, evolution by natural selection is not the smooth diffusion of an advantageous trait, but is mediated by individual births and deaths.

From this perspective, the concept of entropic force has some suggestive power. Entropic forces are forces that originate in the statistical tendency of complex systems with many degrees of freedom to increase their entropy. Natural populations are not in thermodynamic equilibrium, and hence the concept of thermodynamic entropy has little meaning in this context. However, the analogy does hold insofar as natural selection is an effective population-level force that originates in a tendency of the population to reorganize its structure through interaction with the environment.

Does this suggestion not undo the effort to argue against the epiphenomenalism of selection? Not necessarily, because even though entropic forces emerge out of lower-level interactions and thus are not fundamental, they can work in displacing objects. Osmosis is an example of a process driven by an entropic force. They have real effects, can be manipulated, and can be ascribed a direction and magnitude. Entropic forces meet the philosophical criteria of difference-making causality, whether Woodward's manipulationism, or a probabilistic counterfactual account.

Thus the fact that an entropic force is not fundamental but emerges out of lower-level interactions does not in itself constitute an argument that such forces are epiphenomenal and non-causal. Interestingly, it may be pointed out that a strict distinction between a force being 'fundamental' and a 'statistical byproduct' is increasingly under pressure in contemporary physics. Numerous proposals have been made as to how fundamental forces, including gravity, may actually be entropic in nature (Freund 2010; Verlinde 2011).

6 Conclusion

The complexity of the organism-environment interaction means that the resulting evolutionary change can, in principle, be statistically apportioned in many ways. While statisticalists have argued that this multiplicity implies that any single apportioning does not 'cut nature at the joints' and is instead dependent on arbitrary explanatory interests, I have argued that, at least in some cases, a privileged apportioning can be made. In cases where the population tends towards stable equilibrium, a clear distinction can be made between individual-level interactions that do and do not make a difference to the outcome equilibrium state.

Cases where there is no such tendency are more difficult to judge, but when populations evolving by natural selection tend towards stable equilibrium, natural selection has the character of an entropic force: a non-fundamental force, arising out of statistical tendencies at the level of individual interactions, with magnitude and direction.

Appendix: Instability of H-W equilibria

First, let us investigate when two distributions (p, q, r) and (p', q', r') will give rise to the same H-W equilibrium. Then the following three identities must hold:

$$\begin{aligned} p + q + r = p' + q' + r' = 1 \\ p + q/2 = p' + q'/2 \\ r + q/2 = r' + q'/2 \end{aligned}$$

These equations are dependent, and taking r' as a parameter, we get the following set of solutions $\{(p-r+r', 1-p+r-2r', r') | r' \in [0,1]\}$. This can be simplified with the change of variable $\delta = r'-r$, and thus we can say that the basin of the H-W equilibrium $((p+q/2)^2, 2(p+q/2)(r+q/2), (r+q/2)^2)$ is

$$\{(p+\delta, q-2\delta, r+\delta) | \delta \in [-r, 1-r]\}.$$

The basin of a single H-W equilibrium point is the line with direction $(1, -2, 1)$ in distribution space. As one would expect, by letting $\delta = (p +$

$q/2)^2 - p$ one can see that the H-W equilibrium itself is part of its own basin.

Hence we may conclude that there is no open neighbourhood $\mathcal{N}$ around any H-W point (p, q, r) such that N is enclosed by the basin of (p, q, r). H-W equilibria are stable only along one specific line, and hence are unstable equilibria.

Chapter 6

A Global Symmetry in Evolution by Natural Selection

To recapitulate the challenge of causal complexity in the context of the Modern Synthesis: there is no reason to expect that natural selection should cause any non-contingent directional trend. Even if a directional trend occurs, it is contingent on specific environmental states, due to what we called the 'local' nature of natural selection: species can be locally adapted only to a specific environment, and what is adaptive in one environment may not be in the next environment. Natural selection does not look ahead, nor is it in any way a teleological process. It merely describes how different organisms reproduce at different rates due to certain traits, and as described in the previous chapter, this process is dependent on the features of the immediate environment. If the immediate environment changes, so does the direction of natural selection. In this way, the causal structure of natural selection implies an interpretation of evolutionary history as a succession of adaptations, without any large-scale directionality that could characterize all possible histories.[1] Put differently yet again, natural selection may allow for local symmetries towards outcomes, but not for a global symmetry towards an outcome.

This conclusion about evolutionary history, together with the underlying understanding of the basic concepts of environment and natural se-

[1] 'Large-scale' refers to the temporal scale characterizing evolutionary history as a whole. Also referred to as 'global'.

lection, has been challenged by many different evolutionary biologists. It has been argued that there are certain privileged adaptations that are adaptive across a broad range of environment-types. Candidates for such adaptations, which I term *general adaptations*, range from body size (Bonner 1988) and energy-intensiveness (Vermeij 1987), to autonomy (Rosslenbroich 2006), information sensitivity (Ayala 1988; Simpson 1944a) and functional complexity (Bonner 1988). Such adaptations are used as *measures* – degree of size, energy-intensiveness, complexity, and so on – to define corresponding *trends* driven by natural selection.

Because the causal mechanisms underlying the selection for general adaptations are common across all or almost all environments, trends in general adaptations are often presented as not being contingent on the precise succession of environments in evolutionary history.[2] In the terminology of this dissertation, trends in general adaptations can be seen as large-scale features of all (or almost all) possible evolutionary histories, despite variability in environments – precisely because selection for these general adaptations is constant or nearly constant across all possible evolutionary trajectories.

We can locate the point of conflict more precisely by restating the argument against non-contingent trends:

(P1) Static environmental states determine the outcomes of evolution by natural selection;

(P2) Environment-types are contingent occurrences in evolutionary history;

∴ (C) Any possible large-scale trends are contingent occurrences in evolutionary history.

By arguing that certain types of environment are much more frequent than others (i.e., the environments that favour general adaptations), the argument for general adaptations challenges (P2).

In this chapter I will argue that this strategy fails to establish any non-contingent trends, mainly because a trend will be broken if the environment changes towards extreme, hostile conditions – regardless of how frequent the favourable environment is. A high frequency does not guarantee that

[2]Most biologists (except Van Valen, who is unapologetic in calling the Red Queen dynamic and constant extinction rates an evolutionary 'law'; see Valen 1973: 16) avoid the loaded term 'law'. Nonetheless, their arguments strongly suggest that they consider the generalizations they propose to be non-contingent.

a trend in general adaptations may be expected; one would also need to assume that such extreme and hostile environments never occur.

Instead, I will draw on recent developments in the Extended Synthesis to challenge (P1). Not only does variability in states lead to selection for some outcomes (phenotypic plasticity), but outcomes also may influence the environment and resulting selection pressures (through niche construction). The causal relation between environment and outcome of evolution by natural selection resembles something of a causal feedback structure, and I will use this to argue that, even if environment-types are contingent, certain outcomes (specifically, a trend in phenotypic plasticity) need not be contingent.

Central to this argument will be the non-contingent of variability in environments. Recall how in chapter 5 we described the environment of an organism as characterized by many degrees of freedom, and, more often than not, highly variable. Variability is the rule, not the exception. Adding biotic interactions to these environments only serves to increase complexity and variability. From this characterization of the environment, one may conclude that while any specific environmental state may be contingent, *variability* in these states is not a contingent occurrence.

However, one cannot simply conclude from this that a trend in adaptations to variability is to be expected given the nature of natural selection. In particular, there are two main obstacles that this argument would need to overcome in order to be successful. The first is to show how plasticity can be thought of as something that can accumulate. As we will see, plasticity is often quantified as the slope of the reaction norm. This definition reflects an understanding of plasticity as a property specific to a trait of an individual organism in relation to a specific environmental variable (following Bradshaw 1965: 122). According to this definition, it makes no sense to speak of the plasticity of a lineage; there is only a plasticity of a trait in relation to an environmental variable, and plasticity increases when equal amounts of change in the environment produce larger changes in the trait.

By contrast, I will draw on a rival concept, also identified by Bradshaw as 'phenotypic flexibility', that is a property of the whole organism. I will simply call this plasticity as well (it remains unclear to what extent phenotypic flexibility is not simply a form of plasticity: see Nicoglou 2015),

and will identify it with the number of environmental variables to which an organism reacts. Thus, the plasticity of an organism (or species) may increase if the organism becomes able to pick up cues in the environment it was previously insensitive to, and thus in modulating its phenotype, integrates more information about the state of the environment.

The second challenge is to show that increases in plasticity are robust against the occurrence of suboptimal or hostile environments. Insofar as plastic responses are adaptations (some plastic responses may also be the default state of organisms), they are adaptations to certain *patterns of variability*. There is no form of plasticity that is adaptive to variability as such. This is a problem for the argument for a non-contingent trend in plasticity, since these patterns of variability are contingent occurrences in evolution (unlike variability itself).

However, I will argue that some increases in plasticity are ratchet-like. Most importantly, many instances of plasticity, since they represent sensitivity to environment variables, are coupled with various forms of niche construction, where the organism may modify the environment so as to enhance fitness. After a lineage undergoes a plasticity increase in favourable environments, its decrease in subsequent unfavourable environments will be made less probable by niche construction. In this way, the causal feedback structure between organism and environment implied by plasticity and niche construction is crucial for arguing that the contingency in patterns of variability does not imply contingency in a trend in plasticity.

It is important to emphasize what the target of this argument is. It is not to show that actual evolutionary history is necessarily characterized by a trend in increased plasticity. It is merely to show that natural selection, together with the complexity and variability of the environment, implies an *a priori* expectation for such a trend in increased plasticity. Such an expectation can be confounded in specific evolutionary histories, since evolutionary history as a whole is influenced by many causal processes other than natural selection. Evolution by natural selection is an idealized evolutionary history, and for such an idealized evolutionary history, a trend in plasticity is not a contingent pattern.

The chapter is structured as follows. In the first section I elaborate on the causal feedback structure of natural selection, and expand on the phenomena of phenotypic plasticity and niche construction. In the second

section I outline the argument from general adaptations, and in the third and fourth sections I examine how a macroevolutionary trend in plasticity may be understood.

1 Background: Natural Selection in the Extended Synthesis

There have long been dissenting views to the Modern Synthesis; however, developments starting in the 1980s have increasingly led to calls to expand the original framework (Carroll 2000; Love 2003; Müller 2007; Pigliucci 2007). Evolutionary developmental biology (EvoDevo) has played a significant role in putting pressure on the view that evolution can simply be characterized in terms of changes in gene frequencies. As already discussed, developmental constraints may affect further evolution (Smith et al. 1985). Developmental processes may have other effects on evolution: they may facilitate evolution through increasing evolvability, or increase phenotypic variation in response to environmental variation through plasticity. In this way, an important goal of EvoDevo is to map how developmental processes can affect evolution.[3]

As mentioned in chapter 4, how precisely the Extended Synthesis should be defined is still controversial – and indeed, it is still unclear to what extent the theories of the Extended Synthesis can fit within the theoretical framework of the Modern Synthesis. For purposes here, we can simply target an externalist view of evolution by natural selection, where the structure of the external environment is thought to be sufficient to understand evolution by natural selection. Such a view is reflected in Mayr's distinction between 'proximate' and 'ultimate' causation in evolution. The proximate causes for evolutionary outcomes are developmental programmes; however, to understand why such developmental programmes evolved in the first place, we need to take into consideration the ultimate causes, which are, for Mayr, selective pressures.

Whether to attribute this externalist framework to the Modern Synthesis remains a point of continuing contention (see chapter 4, or, for example,

[3]Another important goal of EvoDevo is to explain how developmental processes evolved in the first place, and how they changed over time. See Love 2015.

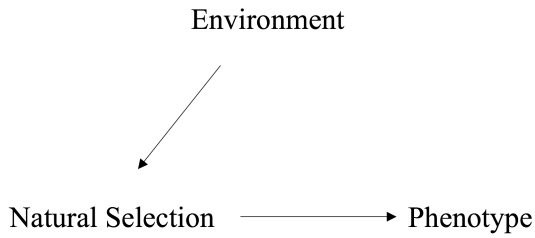

Figure 6.1: Unidirectional representation of the causal structure of natural selection.

Dickins and Barton 2013). I will follow Laland et al. (2013) in doing so, partially because Futuyma's five tenets characterizing the Modern Synthesis do not say anything about the dynamic structure of the environment, and partially because even the proponents of the hegemony of the Modern Synthesis recognize that the macroevolutionary significance of plasticity and niche construction remains underemphasized in the Modern Synthesis. However, it is not crucial to the argument; more important is the way reciprocal causation between organism and environment may imply a large-scale evolutionary trend.

Consider the unidirectional structure of natural selection in Figure 1.Note that the causal structure of natural selection is more complicated than represented in the figure above (see chapter 5). Strictly speaking, natural selection does not cause any novel phenotypes to emerge, and thus does not 'shape' phenotype to 'match' features in the environment. There may be genetic constraints or developmental constraints preventing the optimal phenotype from appearing in a population. However, since in this chapter we are considering evolution by natural selection in an idealized setting, we can ignore such constraints, and follow biologists (such as Laland et al. 2013) in speaking of how natural selection 'shapes' organisms.

Laland et al. (2013) argue that the phenomena of plasticity and niche construction show that how it is inadequate to work with this unidirectional causal model. Phenotypic plasticity shows how the environment can directly shape phenotype, without the action of natural selection, through affecting development, behaviour or physiology; niche construction shows how phenotype can affect environment by modifying features in the exter-

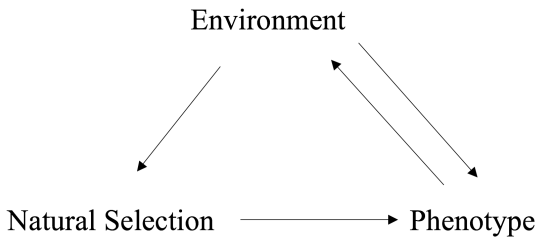

Figure 6.2: Causal feedback occurs when plasticity and niche construction are taken into account.

nal environment, so that the magnitude and direction of natural selection are modified as well.

Before examining the consequences this causal structure has for macroevolutionary trends, I will first give a broad overview of the various types of plasticity and niche construction.

1.1 Phenotypic Plasticity

1.1.1 ANOVA and reaction norms

Phenotypic plasticity, in the broadest understanding of the term, refers to all environmentally induced phenotypic variation. This means that a single genotype can exhibit multiple phenotypes according to the state of the environment. Traditionally (following Scheiner and Lyman 1989), phenotypic variation is expressed through Analysis of Variance techniques (ANOVA[4]) as:

$$\sigma_P^2 = \sigma_G^2 + \sigma_E^2 + \sigma_{G\times E}^2 + \sigma_{\text{err}}^2$$

where the total variance in the phenotype (σ_P^2) is partitioned into three subvariances: variance in phenotype due to differences in genotype alone (σ_G^2), variance in phenotype due to differences in environment alone (σ_E^2) and variance in phenotype due to differences in how genotypes interact with the environment ($\sigma_{G\times E}^2$). The error term σ_{err}^2 captures the effect of any

[4]This statistical technique, originally developed by R.A. Fisher to analyze phenotypic variance, but now widely used across all domains, seeks to attribute the variation of one variable (P) to variation in several other variables (G, E, $G \times E$).

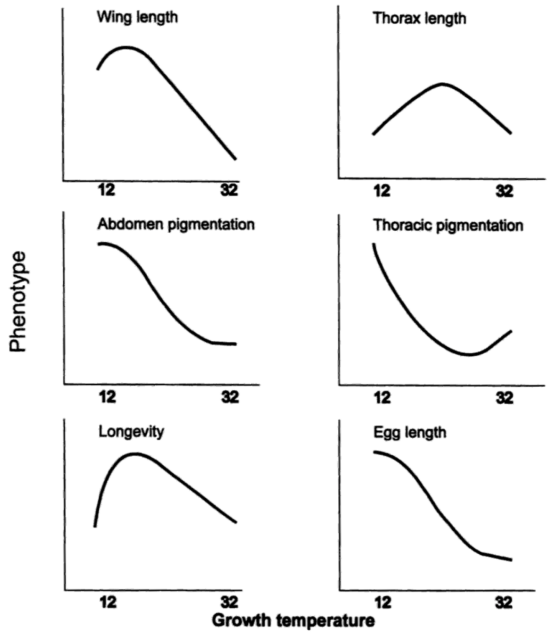

Figure 6.3: Various nonlinear reaction norms in Drosphilia according to growth temperature (Reproduced with permission from David et al. 2004).

process that may affect the measured phenotype but is not attributable to the three previous factors (such as chance fluctuations in otherwise identical environments, developmental errors or even experimental errors).

The equation above has a visual interpretation in the context of reaction norms. Originally proposed by Woltereck (1909), a reaction norm is the function that maps an environmental variable (e.g. temperature, salinity, etc.) onto a phenotypic variable for a given genotype (Scheiner 1993; Schlichting and Pigliucci 1998; Chevin and G. M. Lande 2010). In practice most reaction norms will be nonlinear (Figure 6.3).

In this way, the reaction norm is a simplified and quantitative representation of the developmental processes intermediating the environment and the expressed phenotype. When the reaction norm is flat, then a change in the environment does not lead to any change in phenotype, and hence plasticity is zero. Often plasticity is modeled as a linear reaction norm,

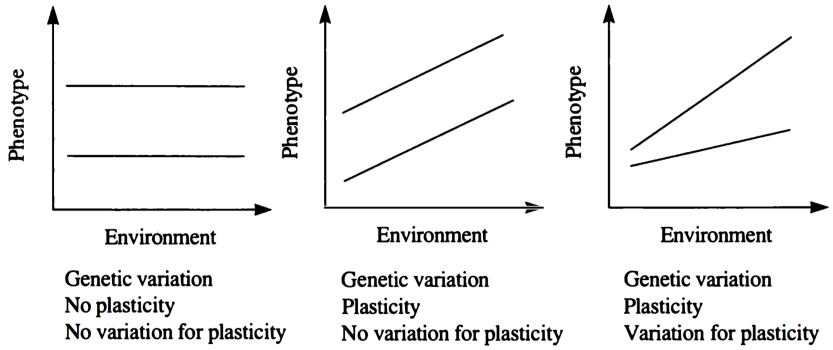

Figure 6.4: Illustration of how genetic variation, plasticity, and genetic variation in plasticity come apart. (Reproduced with permission from Pigliucci 2001)

where the change in phenotype is proportional to the change in the environmental variable. In this case the degree of plasticity can be represented by the slope of the linear reaction norm.

Linear reaction norms allow for a straightforward interpretation of the ANOVA analysis (Pigliucci 2001). In that case, σ^2_E is nonzero if and only if the reaction norm is not horizontal (i.e., the degree of plasticity is nonzero). When σ^2_G is zero, then the reaction norms of different genotypes are identical; otherwise it quantifies the average difference in height between reaction norms of different genotypes. Finally, when $\sigma^2_{G\times E}$ is zero the reaction norms of different genotypes are parallel (indicating equal plasticity); when nonzero, the different reaction norms have different slopes. In other words, the degree of plasticity varies with genotype.

1.1.2 Types of plasticity

This understanding of plasticity, while abstract, is also very broad and corresponds to a similarly broad range of phenomena. One classic example is how the stems of some plants (such as dandelions) elongate when there are other plants nearby competing for sunlight. Another is how some animals develop defenses in response to the presence of predators (such as snails growing thicker shells in presence of crab predators: Trussell and Smith 2000), while others will, sometimes in response, develop more

powerful ways of breaking through these defenses (crabs grow larger and more powerful claws when exposed to armored prey: Smith and Palmer 1994).

This richness in the phenomena of plasticity involves very many types of processes, which are not directly captured in the abstract characterization above. One way to systematize the phenomena is to categorize them according to *environmental source*. Environmental sources may be further subdivided into abiotic and biotic variability. Examples of abiotic variability include changes in sunlight, oxygen, temperature, humidity, salinity, quantity and type of nutritive sources (see Whitman and Agrawal 2009 for an overview). Examples of biotic variability include the presence of predators, parasites or competitors. Note that, at a more fundamental level, biotic variability may be reduced to a special case of abiotic variability, since the proximate cause of an organism's response to predators or parasites will be changes in its abiotic environment (such as changes in sunlight, or changes in chemical gradient).

A second dimension along which types of plasticity are categorized is the type of trait affected. Thus plasticity can occur in biochemistry, physiology, morphology, life history and behaviour (Whitman and Agrawal 2009), and in each case very different causal processes are involved. Some changes, such as changes in life history (timing of reproduction or maturation), are irreversible, and involve a host of developmental processes. Such changes in life history often affect many other traits, from hormone concentrations to body size. Other changes, such as in physiology (colour, metabolic rate, digestion, etc.), may occur in a relatively isolated fashion and may be reversible.

An important distinction is between *developmental conversion* and *phenotypic modulation*. Developmental conversion refers to plastic changes that are irreversible and discrete, such as the choice of sex or of social caste in some species; phenotypic modulation refers to plastic changes that are reversible and continuous, such as body size. These two types of plasticity represent extremes (Whitman and Agrawal 2009): most real instances of plastic change will be capable of at least partial reversion, and will involve some spectrum of intermediate forms.

Because plasticity may also refer to changes in the biochemistry of an organism (e.g. enzyme or hormone concentrations), another important dis-

tinction is between *active* and *passive plasticity*. Passive plasticity refers to a more or less automatic 'susceptibility' of an organism that is not the result of any regulated process. Examples of such plasticity include the ways in which toxins, poor nutrition, extreme temperatures, acidity, oxygen levels and salinity may affect a range of processes, ranging from the chemical and enzymatic to the cellular and developmental. A typical example here is how small size results from poor nutrition. Active plasticity, by contrast, refers to a highly regulated response to environmental variability, where a host of processes coordinate to produce a particular phenotype in response to the environment.[5]

This distinction is closely related to the distinction between *anticipatory* and *responsive phenotypic plasticity*. In anticipatory plasticity, the organism is able to detect certain environmental cues that are indicative of a change in the environment that is beneficial or harmful to the organism. Such cues set in motion a coordinated response; if the cue is undetected, responsive plasticity may lead to an alternative change of phenotype. For example, the detection of a predator may provoke a flight response in an organism; however, if the predator remains undetected, this may also lead to a merely responsive change in phenotype (for example, destruction of the body).

There are some other factors, such as time-lag between cue and response,that characterize a plastic response (see Whitman and Agrawal 2009). We will discuss these in some further detail when considering the environmental conditions of selection for plasticity; however, more important here is to note two difficulties.

First, what should count as a phenotype is not always clear. As DeWitt and Scheiner (2004) remark, plasticity could also potentially refer to changes in phenotype such as the buildup of muscles after use, or to changes in behaviour through learning. Some biologists include only phenotypes that are affected by development within plasticity; others are more liberal and include all forms of learning and behaviour. In this chapter we will follow the liberal use of the term, and include any environmentally induced change in phenotype that may affect the fitness of the organism.

[5]The existence of passive plasticity has led some to claim that plasticity, not canalization, describes the default state of an organism. This will be important later in this chapter.

Another issue concerns how to distinguish active from passive plasticities, since every phenotype is affected by the environment in some way. As noted above, some degree of passive plasticity is the default state of any trait, and is simply a consequence of physical and chemical causal processes affecting an organism. One way to make sense of this is that a trait can be simultaneously characterized by both active and passive plasticity (Whitman and Agrawal 2009: 18): a trait may have an active response to one environmental variable (in the sense of a coordinate response), but only a passive response to another.

1.1.3 Plasticity as a Source of Phenotypic Variation

In the Modern Synthesis, the most important source of phenotypic variation was thought to be genetic variation. Environmental effects on phenotype were recognized, but were not thought to be relevant for understanding how evolutionary trajectories unfolded over time. Often such environmental effects were categorized as developmental noise (Canfield et al. 2009): often quoted in this regard is Fisher's remark "It is not surprising that such elaborate machinery should sometimes go wrong" (Wigglesworth 1961, cited in West-Eberhard 2003). Thus for example, a small size resulting from a lack of nutrition was characterized as a development gone wrong – not as the effect of an independent property (plasticity) that itself could be the object of selection. Plasticity adds an important new source of phenotypic variation, by introducing phenotypes that otherwise might not be expressed. Further, such phenotypes can be inherited, not directly, but through inheritance of the reaction norm. In this way, an environmentally-induced phenotype may spread through a population by natural selection through selection for the underlying reaction norm.

There is still some debate as to how precisely plasticity may affect evolutionary trajectories. Some have suggested that phenotypic plasticity may slow down the rate of genetic change, because a fixed reaction norm would remain viable for longer than a fixed phenotype (Price et al. 2003).[6] The threshold at which environmental change will favor genetic change is higher when the organism has some degree of plasticity at its disposal. In

[6]Note that this is a form of counteractive niche construction: while the actual physical environment is not modified, the selection pressures are.

this way, genetic change would be expected to occur faster in the absence of plasticity.

Others have suggested that the opposite may happen: phenotypic plasticity may speed up the process of genetic change. Drawing on the Baldwin effect (1896) and the concept of genetic canalization (Waddington 1953), West-Eberhard sketches a process how this may happen (West-Eberhard 2003). First, plasticity produces novel phenotypes in response to evolutionary change. However, as the environment settles down into a more static state, there will be selection for individuals in a population who are able to produce the novel phenotype without cue from the environment (since the latter can lead to detection errors, such as false positives and negatives). This selection leads to a canalization of the trait in a process of 'genetic assimilation'. In this scenario, plasticity is the 'leader' of evolutionary change, and genes are the 'followers'.

In sum, plasticity represents a way in which the environment may play a dual role in evolutionary by natural selection. Not only does the environment select genetically produced phenotypic phenotypes (through natural selection), but it may also directly influence phenotype through plasticity. This will then indirectly change the nature of selection, as the pool of phenotypic variation is changed.

1.2 Niche Construction

Niche construction singles out a type of phenotype that is of special interest to the process of evolution by natural selection, namely behaviours of organisms that modify their selective environment. In the same way that plasticity brings into focus the dual role of the environment in evolution, so niche construction highlights the dual role of the phenotype (or organism): first as object of natural selection, whereby the genotypes coding' for various phenotypes is differentially transmitted to the next generation, and second as cause of change in the environment of the organism, and thus of change in the magnitude and direction of natural selection. An influential general theoretical case for niche construction was made by Lewontin (1983; 2000), who formulated it in terms of a pair of differential equations:

$$\begin{aligned} \frac{\mathrm{d}O}{\mathrm{d}t} &= f(O,E) \\ \frac{\mathrm{d}E}{\mathrm{d}t} &= g(O,E) \end{aligned}$$

These equations are very general representations of evolutionary dynamics (and best interpreted as a heuristic). The first equation tells us that the way in which organisms (populations) change over time is determined by the properties of the organisms and of the environment. Similarly, the way in which an environment evolves over time is not only dependent on environmental variables, but also on organismic properties. Evolutionary dynamics is considerably more complex than can be captured in this way (see next chapter for a discussion of the statistical nature of natural selection); however, it is a good illustration of the general idea behind niche construction.

We can mention three classic examples of niche construction. The first is how cyanobacteria modified the atmosphere of the Earth, converting it from a reducing environment to an oxidizing one through the production of oxygen (the 'Great Oxygenation Event'). This then was thought to have caused a mass extinction of anaerobic organisms and a strong constraint on further evolutionary trajectories (Irwin and Schulze-Makuch 2010). The second is the way in which earth worms, by burrowing, affect the amount of topsoil and concentration of nutrients in the soil. This change in the environment then affects subsequent selection for the structure of the epidermis and quantity of secreted mucous (Laland and Sterelny 2006). Finally, beavers' dam-building activity has been a key example of niche construction. Dawkins (1982) had originally argued that dam-building was an 'extended' phenotypic trait. In Dawkins's gene-centered view of life, the only genuine unit of selection is the gene, largely since genes, unlike phenotypes, are the only units capable of high-fidelity replication, and so are the only units that meet the heritability criterion for evolution by natural selection. Hence, according to this view, any physical process that is produced by the genotype is considered a phenotype, including beaver dams.

However, according to niche construction theorists, categorizing the beaver dam simply as an extended phenotype misrepresents the causal

interaction between phenotype and environment. In view of Dawkins, beavers build dams because there in the beavers' past there was selection for dam-building alleles over alternative alleles. Niche construction theorists claim that this ignores how how dam-building phenotypes alters magnitude and direction of selection on subsequent generations – in particular, the selective feedback that occurs between initial dam-building activity, and further selective reinforcement of the behaviour (Laland and Sterelny 2006, Laland et al. 2013: 131). The spread of dam-building alleles in a population is likely an effect, not a cause of the evolution of dam-building behaviour.[7]

Laland et al. (Laland et al. 1996; Laland and O'Brien 2010) distinguish some specific types of evolutionary dynamics that not readily explainable in the extended phenotype model: (1) inertia effects, where the evolutionary response to selection only occurs after a number of generations (2) momentum effects, where evolutionary change continues in the same direction despite the stopping or reversal of selection, and (3) sudden extinctions and fixations. These patterns of evolutionary change (i.e. changes in population composition over time) cannot be readily explained without taking into account the niche construction aspect of the traits involved.

In this way, niche construction represents a causal relation between phenotype (cause) and environmental state (effect) and, seen in this way, is the converse relation of phenotypic plasticity. (Note that niche construction traits may themselves be produced as plastic responses to environmental change: this will be important later in this chapter.) Further, just as different types of plasticity may be distinguished according to the details involved in the causal process, Odling-Smee et al. (2003) categorize different types of niche construction according to two axes.

The first axis distinguishes between *perturbation* and *relocation* niche construction. In relocation, the selective environment of an organism is changed by an organism migrating to a novel physical environment. By contrast, in perturbation niche construction, the selective environment is actively changed, by modifying the environmental state of the organism's current physical environment. In each case the selective environment is

[7]For instance, Wells (2015) claims that Dawkins's argument cannot explain how sometimes the host's phenotype is what makes it possible for genes to exist (not the other way around).

changed, but the difference lies in whether the organism actively modifies a physical environment or merely exposes itself to a novel physical environment.

In many real cases, both perturbation and relocation are involved. Relocation will in practice often result in slightly different selection pressures from preceding ones, which then might be responded to by actively modifying the novel physical environment to some degree. And conversely, perturbation may be coupled with some relocation, for example when animals construct nests and burrows but also relocate in order to select a beneficial location for their nests or burrows (Odling-Smee et al. 2003: 45).

The second dimension along which niche construction can be categorized refers to whether the niche construction activity restores an earlier environmental state or creates a new one. The first is called *counteractive* niche construction, and can occur both through perturbation (such as the thermoregulation of nests) or through relocation (such as the seasonal migration of birds). Counteractive niche construction is not necessarily adaptive, but when adaptive it is often because it is fitness-enhancing for the organism to counteract environmental variability and maintain a more or less constant environment. However, such niche construction can also serve to counteract a new (fixed) environmental state by restoring a previous one.

At the opposite extreme is *inceptive* niche construction, by which the environment is modified to a new state. Such inception can also occur by perturbation (by an organism polluting its environment) as well as by relocation (an organism invading a new habitat). Again, such niche construction need not be adaptive, and to emphasize this Odling-Smee et al. (2003) distinguish between *positive* (fitness enhancing) and *negative* (fitness decreasing) niche construction. Invading a new ecosystem, where a population would be at a competitive advantage, would be an instance of positive niche construction, while the depletion of a food source (such as occurs in predator-prey Lotka-Volterra models) would count as negative niche construction.

As with phenotypic plasticity, we mention in passing a key conceptual problem facing the definition of niche construction. First, niche construction may be the default state of an organism, since any organism will have

a causal effect on its environment, if nothing else through its metabolism which requires the consumption of nutritive sources in the environment and the excretion of waste. Second, some biotic interaction with competitors, prey or predators is nearly universal among living species. Hence niche construction may be an exceptionally broad category, including most if not all phenotypes and not necessarily referring to a special category of phenotypes.

We need not take a stand here on issues concerning the delineation of the boundaries of niche construction phenotypes, or plastic phenotypes for that matter, since for present purposes we are primarily interested in how the phenomena of niche construction and plasticity represent *causal pathways* between environment and phenotype that have remained unrecognized in interpretations of evolutionary history. Of interest is the causal feedback loop that may be instantiated (see Figure 3) through phenotypic plasticity and niche construction.

2 Natural Selection and General Adaptations

As mentioned in the introduction, given the dependence of the direction of natural on the specific environmental state, there would seem to be no *a priori* reason why natural selection should consistently favour one trait over the other. If natural selection were to be represented as some force vector, one would have no reason to expect the vectors instantiated in evolutionary history to be pointing on average in one direction rather than another.

This argument can be challenged by pointing to what seems to be a considerable overlap in the causal processes of different possible environments. The challenge can be roughly constructed as follows. Organisms must survive and reproduce in all environments, so if an adaptation were to improve 'general abilities' such as foraging, predator avoidance, fecundity, mating success, tolerance to stress, etc., it would seem likely that such an adaptation would be favoured in most if not all environments. It is always better to be more fecund than less — other factors being equal. Put more abstractly, if an adaptation can take advantage of the overlap of the causal structures of the different environments, and respond to the features that many environments have in common, it will be *generally adaptive*.

It is not even necessary that every single environment must participate in the overlap, only that a 'sufficient' number of environments overlap so that *on average* the general adaptation will be selected for. The relative frequency with which the favourable type of environment occurs needs only to reach a certain threshold; this threshold need not be a relative frequency of 1.

Size has been argued to have these general adaptive benefits (Bonner 1988). Similarly, functional complexity has been argued to allow for a division of labour between the parts of an organism, and thus an increased energy efficiency of the organism as a whole (also Bonner 1988). Increased energy efficiency is always to be favoured by natural selection (again: other factors being equal).

A further example is Vermeij's argument (1987) for a large-scale trend in energy-intensiveness. Vermeij understands energy-intensiveness as a banner term, one that can be manifested in different ways, including larger size, longer life-span, higher metabolic rate, a larger number of interactions with the environment, and a larger number of functions. In a competition between an individual with higher energy-intensiveness and one with lower energy-intensiveness, the former will have the edge, and thus selection will, on average, favour energy-intensiveness. The condition for the selection of increased energy-intensiveness is that there are sufficient environmental resources available to support higher energy-intensiveness. Vermeij suggests climatic warming, the spread of lowland forest, and expanded shallow marine waters as some factors (Vermeij 1987: 377). These conditions notwithstanding, Vermeij characterizes the fitness increase brought on by increased energy-intensiveness as a "nearly universal property" in competitive environments (Vermeij 1999).[8]

There are two problems with this line of argument. The first is that it is unclear why environments should be such that there will be selection on average for general adaptations. This problem can be broken down into two parts: the relative frequency of favourable environments, and the relative weight that environments have for estimating the average selection.

Concerning relative frequency, when we look in detail at the causal

[8]In this way, while Vermeij does not use the term 'law' or 'noncontingent generalization', it plausible that this is what he has in mind in formulating the trend in energy-intensiveness.

mechanisms underlying selection for general adaptations, we see a complex web of conflicting mechanisms. For example, with regard to size, numerous costs have been postulated, including increased foraging risk (Dibattista et al. 2007; Carlson et al. 2008), structural problems, a greater danger of reduced locomotory performance (Lankford et al. 2001), and a decreased potential for adaptive evolution (Dombroskie and Aarssen 2010). The benefits of a general adaptation may be widespread, but so may the costs, and in this way there may be many environments where the costs of the general adaptation outweigh the benefits.

Even if we were to grant the relatively high frequency of favourable environments, this still does not guarantee a selection for the general adaptation on average. For that to occur, selection for the general adaptation would need to outweigh selection against the adaptation, and it is also unclear why that should be the case. For example, with body size, it has been argued that there is selection against size during deteriorating environmental conditions (such as mass extinctions), as larger organisms will be less able to find sufficient nutritional resources in such conditions (LaBarbera 1986; Arnold et al. 1995; Alroy 1998). It has also been noted that in unstable and unpredictable environments there is selection for an increased number of offspring and shorter generation times (r-selection), which are correlated with decreased body size (Bonner 1988: 48). The question then arises as to how to weigh the prevalence of resource-rich and resource-poor environments, or stable and unstable environments? There is no reason to believe that, in the grand scheme of things, the causal mechanisms leading to an increase of some general measure M should systematically trump the causal mechanisms leading to a decrease in M.

The second problem is that, even if there were an average selection for the general adaptation, a trend in the general adaptation would still not be guaranteed. For example, if the selection against the general adaptation were infrequent but severe, the trend might simply be broken. Or, if one type of environment were to be lethal to the general adaptation, any incipient trend would be broken if that environment-type happened to occur. Being adaptive on average does not imply that the increase in some general adaptation would be robust against suboptimal environments.

The upshot of these difficulties is that additional hypotheses would need to be made about the succession of environments in evolutionary history to

guarantee a trend in the general adaptation. A global macroevolutionary trend would require the environments to line up in a particularly beneficial way. Trends in general adaptations are contingent generalizations about evolutionary history, contingent on certain environmental conditions.[9]

Empirical studies have shown little support for the argument that trends in general adaptations even characterize *actual* evolutionary history. While the proposed globally driven trend in size has received some support as a *local* trend in lineages' invertebrates (Chown and Gaston 2010; Lamsdell and Braddy 2010), plants (Chaloner and Sheerin 1979) and vertebrates (e.g. Lamsdell and Braddy 2010), there is much less support for it being a global trend, characterizing evolutionary history as a whole. The comprehensive review by Jablonski (1996) shows that many instances of increases in body size depend on specific features of the environment. A landmark paper by Payne et al. (2009) documents an increase in maximum body size over evolutionary history, but does not note any evidence that the trend is driven by natural selection. Similarly, the increase of functional complexity has received scant support as a trend driven by natural selection (McShea 1996). This is not surprising given the various ways in which selection *against* functional complexity may be much more frequent than selection for (Gould 1996; McShea 2005).

In summary, the theoretical argument for natural selection driving a trend towards the increased development of some general adaptation (increased size, increased complexity, etc.) fails because it is dependent on contingent assumptions about which environments happen to occur. Given what we know about the complexity of environmental structure, there is little reason to believe that environments would necessarily line up in just the way demanded by trends in general adaptations.

If such conditions are integrated into the generalization, the resulting trend at most describes a local symmetry: a trend that characterizes some instances of evolution by natural selection. However, it does not characterize evolution by natural selection generally. Given the nature of environmental structure and of natural selection, there is no reason to

[9]Could one not apply the *ceteris paribus* strategy here, and obtain a noncontingent generalization? Yes, but the generalizations would be vacuously true in some evolutionary histories. In this way, a trend in general adaptation may represent a local symmetry, but not a global symmetry.

expect a global trend in any type of general adaptation.

3 Natural Selection and Variability Adaptations

In the first section I gave an overview of different types of plasticity, and of some of the way in which plasticity may be significant for the dynamics of trajectories in evolutionary history. We passed over the difficulty that different types of plasticity are selected for under different conditions, and that some patterns of variability may lead to no selection or even selection against plasticity. In the current section I will outline these selective conditions in some more detail.

3.1 Sources of Selection for Plasticity

Under what conditions will there be selection for plasticity? A first estimation is given by the *tolerance curve*, which represents the value of fitness that can be expected in a given environment. The tolerance curve for a non-plastic individual (see Figure 3) is a reflection of how adaptable a given phenotype is in a particular environment. By contrast, the tolerance curve for a plastic individual shows what fitness is to be expected as the individual changes its phenotype in response to the environment. The effect of plasticity is to flatten the tolerance curve (lower maximum, but larger variance), so that the geometric mean fitness will be raised when there is environmental variability.

Besides environmental variation, another important condition determining whether plasticity is selected for is whether the plastic expression of traits tracks the state of the environment with sufficient accuracy (Moran 1992; Sober 1994; Godfrey-Smith 1996). One specific way in this condition can be violated is if the environment lacks a reliable cue. Another way is if the time scale of the environmental variation is too extreme. The reason for this is that, assuming there to be some time-lag between the change of environment and the change of trait, there will be a brief moment in time when the trait of the organism does not track the environment. When the tracking mechanism of the organism can no longer keep up with the frequency of environmental variation, a rigid response will be favourable.

In general, the specific type of *pattern* of variability will play a crucial

role in determining what type of plasticity is favoured by selection — or if plasticity is to be favoured at all. Some important parameters that define a pattern of variability are (1) the time scale of temporal fluctuation, (2) the scale of spatial fluctuation, and (3) the presence or not of a cue to detect oncoming environmental change. Figure 6.5 (adapted from Pigliucci 2001) shows various permutations of these selective conditions together with the evolutionary outcome that may be expected.

Pattern of Temporal variation	**Pattern of Spatial variation**	**Expected Evolutionary Outcome**
Scale: Shorter than generation time	/	Physiological Plasticity
Scale: Longer than generation time Distribution of variation: random	/	Adaptive coin flipping Genetic polymorphism
Scale: longer than generation time Distribution of variation: regular	/	Plasticity by seasonal or generational forms
/	Scale: coarse-grained Distribution of variation: random	Adaptive coin flipping Genetic polymorphism
/	Scale: fine-grained Continuous phenotype	Phenotypic plasticity by modulation
/	Scale: fine-grained Distinct phenotype Time-lag between environmental input and phenotypic response	Phenotypic plasticity by developmental conversion

Figure 6.5: An overview of different patterns of variability, with different expected outcomes.

Note that variability need not always favour plasticity. Some patterns of variation favour non-plastic adaptations to variability, such as genetic polymorphism, which means that many different phenotypes are present in a population. This represents a form of bet-hedging by a population, and occurs whether or not the individuals in the population are plastic with respect to the phenotype.

We need not go into further detail into the conditions under which

plasticity is selected (but see Pigliucci 2001). For the purposes of this chapter it is sufficient to establish that plasticity is not always adaptive even if there is variability in the environment, and if it is adaptive, that there are different *types* of plasticity which are adaptive according to the *pattern* of variability in the environment. This means that if there can be any trend in plasticity, it cannot be a trend across all types of plasticity, nor can it result from a ubiquitous selection for these various types of plasticity.

3.2 Plasticity and Niche Construction

The net effect of plasticity on fitness is sometimes represented as a flattening of the tolerance curve, decreasing fitness variance in fluctuating environments (Figure 3). For this reason, plasticity is sometimes represented as a *buffer* against environmental variation (e.g. Grove 2014; Ellis et al. 2009; Kearney et al. 2009), and thus as a form of counteractive niche construction. Lewontin (1983) suggests that many niche construction phenotypes, such as migration, hoarding, habitat selection and thermoregulation, are adaptive because of this buffering that is made possible by plastic responses to temporal variation in environmental resources (Laland et al. 1996).

This is closely related to the 'retarding' effect plasticity may have on genetic change: plasticity in a trait gives an organism a repertoire of possible phenotypes instead of a single genetically determined one, and hence it decreases the chance that the organism will be maladaptive in the succession of the different environments. This increases the average fitness (averaged over a succession of environments) over that of a non-plastic organism, whose fitness will vary directly in accordance with environmental variation.

Further, the evolution of plasticity may also be paired with inceptive niche construction (Odling-Smee et al. 2003; Grove 2011), since the plastic response to variation in the environment may entail habitat expansion and relocation. Relocation, whether by means of migration, foraging or sensing a predator, allows an organism to respond to a changing environment by extricating itself from suboptimal conditions. Thus relocation helps buffer against environments that would otherwise be unfavourable, simply by avoiding these unfavourable environments.

Relocation is also one of the oldest adaptations to environmental variability. Consider undirected motility in bacteria. Undirected motility has been hypothesized to be adaptive in spatially heterogeneous environments because motility allows bacteria to distance themselves from others and thus have access to more resources (Wei et al. 2011). Further, when motility is removed by researchers from bacterial strains, it rapidly reappears, indicating a strong selection for motility (Taylor et al. 2015).

3.3 Distinguishing Selection for Plasticity and for General Adaptations

At this point, it may be helpful to compare the selection for plasticity and the selection for general adaptations. In a sense, plasticity and general adaptations are superficially similar since both result in a higher average fitness across diverse environments. However, the causal process underlying each is very different. A general adaptation is adaptive in virtue of specific causal processes that are common to a large proportion of environments (for example, biotic competition); by contrast, plasticity is not adaptive to any specific causal features of an environment. Plasticity gives higher average fitness in fluctuating environments, simply because it allows an organism to use a phenotypic repertoire as a hedge against environmental change. It does not matter what the specific causal processes are that define the organism-environment interaction, as long as those processes change over time. General adaptations give a higher average fitness by reacting to a constant core in fluctuating environments; plasticity gives higher fitness by reacting to the fluctuation itself.

In this way one can distinguish between variability and generality selection (see also Potts 1998). These are two separate causal processes: the former selecting adaptations to specific patterns of variability (which may or may not be frequent) and the latter selecting adaptations to general features across environments (which are assumed to be frequent). However, their effects cannot always be easily distinguished. The selection for size can lead to increased plasticity, as the increase in body size can lead to a net increase in new behavioural and developmental pathways; conversely, the selection for plasticity can drive an increase in size. It depends on whether size is a response to specific features in an environment, or to variability in the features of an environment.

An implication of this is that the same trait can be seen both as a general adaptation and as a variability adaptation. For example, sex has been hypothesized to be a reaction to environmental variability or to biotic competition (Scheiner 1993).

4 An Intermittent Global Trend

In this section I will argue that the mechanisms of selection for plasticity imply an a priori expectation of a global macroevolutionary trend in plasticity. Two conditions are necessary for a sustained trend, both of which must be argued for. One of these is that plasticity must be a measure that can increase on a macroevolutionary scale — something that does not make sense if plasticity is understood merely as the slope of a reaction norm. The other condition is that a trend in plasticity must be sufficiently robust against periods of nonoptimal environmental variability, during which there may be selection against plasticity. If these conditions are met, then it no longer matters how frequently favourable environments appear, since even infrequently occurring favourable environments ensure a trend if the increase in plasticity is ratchet-like.

4.1 Defining a Plasticity Measure

The most common plasticity measure is the slope of the reaction norm (Bradshaw 1965). As the slope of the reaction norm increases, the same amount of environmental change results in a larger difference in phenotype, and for this reason it is said that the amount of plasticity in the trait is 'larger'.

However, this plasticity measure is not a suitable candidate by which to define a global trend, for two reasons. First, a reaction norm is a specific adaptation to a particular pattern of variability, and will increase or decrease as the environment changes. Second, the slope of the reaction norm is associated with the property of a particular trait of an organism, and cannot be used to compare two different organisms in the succession of different selective environments. Hence, the slope of the reaction norm cannot be used to compare the plasticity of different species' lineages if they are inhabiting different selective environments.

Nonetheless, there is a meaningful extension of the common measure of plasticity to the macroevolutionary context, and one that follows Bradshaw's general idea of 'phenotypic flexibility' (Bradshaw 1965). According to Bradshaw, an organism may have a 'flexible' response in some trait, even if the organism does not change its phenotype in response to variability in the environment. An example of this is *homeostasis*, where a phenotype is maintained in face of environmental change across many different parameters.

Today, homeostasis is often categorized as a type of plasticity (see Whitman and Agrawal 2009: 21), even though it presents difficulties for a definition of plasticity solely in terms of reaction norms. In fact, increased plasticity in one trait (in the sense of an increased slope in the reaction norm) may lead to increased canalization in another trait, and vice versa. The apparent lack of plasticity in one trait can reflect some buffering mechanism in another trait – and this argument is only strengthened when we consider how some plasticity is likely a default state for many if not all traits. In this way, the linkages between these different traits make it difficult to analyze the evolution of plasticity in a trait by isolating the trait from the organism as a whole.

In light of some of these complications, some (e.g. Forsman 2015) have called for a 'whole organism' approach to plasticity: plasticity should also be thought of as a property of an organism as a whole (in relation to the environment), not simply as a property of a trait (in relation to the environment). However, the whole organism approach is related to the trait-specific approach, and Forsman advocates that the trait-specific plasticity of many different traits be analyzed simultaneously. By means of multivariate statistical analysis and composite measures of plasticity, these analyses then can be synthesized to estimate differences in plasticity between organisms (Forsman 2015: 280).

With this theoretical background, we may make use of the following heuristic definition of the degree of plasticity of an organism:

> The **degree of plasticity of an organism** is the number of environmental variables to which an organism has an active, coordinated plastic response.

While this definition may or may not be of empirical use (I will not consider

how such a degree should actually be measured), it does reflect the main idea behind the 'whole organism' approach to plasticity, and, as I will argue, is of heuristic value for considering how plasticity may accumulate in evolution.

One consequence of the definition is that plasticity may increase, while the phenotype remains unchanged. This is the case, for example, when the homeostasis of an organism becomes more robust. In this way, this definition overlaps with the way robustness (understood as an organism's buffering capacity against environmental variability) is measured in biological systems theory (Kitano and Oda 2006). In that theory, the degree of robustness reflects the number of 'input networks' of an organism, which in turn is a measure of the organism's detection capacity of the environment.

Another consequence of this definition is that plasticity may increase simply by migrating to a new environment. Say organism A migrates to a new environment, and A discovers that it has a previously latent plastic response that is activated only in its new environment. Then A's plasticity has increased by changing environment, since in the process it has increased the number of environmental variables to which it is sensitive (i.e., so as to produce an active, coordinated response).

4.2 A Partial Ordering

It is important to note that plasticity thus understood is a *partial ordering*, as different organisms may be sensitive to different environmental variables. Are phytoplankton and redwood trees equally plastic? Both are sensitive to sunlight, but otherwise inhabit very different environments, and the large size difference almost precludes that they would be sensitive to the same environmental variables (for example, phytoplankton are acutely sensitive to gradients of macronutrients while the same gradients might be imperceptible for a redwood). It may not be possible to judge one lineage to be more plastic overall than another lineage.

This has implications for the shape of any possible trend associated with this measure. A total measure means that for any two entities, either one is greater than the other or both are equal, implying that a trend would have a *linear* structure. By contrast, a trend in plasticity, as understood here, would branch out in different mutually incommensurable directions, even while retaining an overall directionality.

That such a directionality is possible, at least in principle, is ensured by the lack of a hard upper bound for the increase of plasticity. The physical size of the selective environment may be bounded, even though it has increased over many orders of magnitude in the course of the evolution from small unicellular organisms to large multicellular ones. However, the number of environmental variables an organism can be sensitive to would not seem to be affected by such boundaries. Any physical environment has an intractable number of causal processes that have an effect on an organism's viability and reproductive capacity; hence in principle, a plastic response to each variable process could be developed (even though size would be an obvious constraint). Further, the creation of socially selective environments adds a new whole new and potentially unlimited source of variability (e.g. through Machiavellian interactions, Humphrey 1976), and hence also additional scope for plasticity to increase.

4.3 Frequency of Selection for Plasticity

In the previous section we discussed the conditions under which the various types of plasticity are selected for. If the scale of temporal or spatial heterogeneity in some environmental variable is too long or too short, plasticity will not be favoured. Another important condition is the detectability of the environmental cue alerting the organism to changed environmental conditions. Further, even when plasticity is favoured, different types of plasticity will be favoured according to the specific nature of the environmental variation (Figure 6.5). However, at this point in the argument, we can abstract away from the details concerning which conditions are and which are not favourable for plasticity, and speak only of 'favourable' and 'unfavourable' selective conditions for plasticity. Unfavourable environments include both those in which there is no selection for plasticity increase, and those in which there is selection for plasticity decrease.

At first sight, it would seem to be important for our argument to take into consideration the frequency with which such optimal environments occur. Any assumption about the frequency of optimal environments would constitute a contingently true proposition about what types of environment occur in evolutionary history. Making an assumption would imply that the resulting trend would also be a contingent pattern in evolutionary history.

To address this, let us simply let p represent the probability of favorable environmental conditions (for plasticity increase). This merely labels our ignorance, since we have no reliable method to estimate what the actual value of p would be: the probability represents an average probability over all possible environments that organisms could inhabit. Given the complexity of environments, this calculation would be intractable. However, we will not assume any particular value of p except that it be nonzero: optimal selective conditions for plasticity occur at least sporadically.

4.4 Relative Frequency is Non-zero

I will now argue that we may assume that all patterns of variability occur with nonzero probability. Variability as such would seem to be necessary for biological evolution. Given that life itself arises in a far-from-equilibrium environment characterized by chemical and temperature gradients (e.g. England 2013), static environments are likely to be incompatible with life and do not need to be considered.

Within this general variability, the multidimensionality suggests that this variability will be organized in various patterns, since the environment of any organism contains a multitude of causal processes, all of which are changing at different rates. This inevitably creates very complex patterns of variability, where different environmental variables are fluctuating at different time scales. For example, precipitation may undergo different fluctuations at different time scales: random at the scale of days, seasonal at the scale of months, uniformly increasing at the scale of centuries (climate change). Or, if the environment includes biotic interactions, such as predator-prey interactions, the variability in the selective environment will have an added layer of complexity. The upshot is that, unless contingent assumptions are made about environmental structure, there is no particular reason why those particular patterns of variability that favour plasticity should be more common than the patterns that do not. No pattern of variability is privileged, so also the selectively optimal patterns of variability should occur sporadically.

4.5 Robust vs. Non-robust Increases in Plasticity

Once contingent assumptions about the structure of the environment are avoided, the greatest challenge is now to argue why increases in plasticity, in contrast to increases in general adaptations, are *robust*. It is necessary to argue this, because after some initial increase in plasticity in a lineage, an unfavourable environment, such as a period of prolonged stasis, is bound to set in at some point. However, there is reason to believe that, because of the unique properties of plasticity, increases in plasticity should be sufficiently robust against unfavourable conditions.

To make progress here, we must analyze some of the general processes by which lineages can lose plasticity in unfavourable environments. I will consider three of the most important: viability selection, interspecific competition (competition with other lineages), and intraspecific competition (in the form of genetic assimilation). It will turn out that it is not possible to argue that *all* types of plasticity are robust: we must assume some form of inceptive niche construction. This represents an important restriction of the conclusions reached here: a trend is to be expected in certain types of plasticity only, that is, those types of plasticity that also have the ability (or increased ability) to avoid hostile environments.

First consider viability selection. As mentioned previously, plasticity entails that an organism has a repertoire of phenotypes at its disposal, and so if the environment fixates at a particular state, plasticity increases the probability that an organism will be able to produce a phenotype more adapted to the external environment than if the organism did not have any plastic response. This is the buffering capacity of phenotypic plasticity, and is a simple consequence of a plastic organism having a repertoire of phenotypes available. A random rigid response, with a narrower tolerance curve, is more likely to miss out on a viable response to a static environment.

However, plasticity does not guarantee survival in the face of environmental change. An obvious challenge to this buffering aspect of plasticity occurs when the environment undergoes extreme change, to a state outside the range of its previous fluctuations. In such a case the plastic lineage may no longer be viable. Thus plasticity merely increases the probability that a lineage will be robust in an unfavourable environment.

Consider the situation when the unfavourable environmental lasts for

a long time. Under such conditions, the viability-increasing property of plasticity is not sufficient to ensure robustness, as one would expect the plastic response to disappear through competition with non-plastic lineages, where either a fixed response, or genetic polymorphism or adaptive coin-flipping is favoured.

In this way, some patterns of variability, if not responded to, would indeed entail a break in the trend in plasticity. Extended stasis would induce genetic assimilation (a form of intraspecific competition, where nonplastic individuals are selected); extreme variability might hand a competitive advantage to groups or lineages characterized by adaptive coin-flipping or genetic polymorphism (Pigliucci 2001).

However, some forms of plasticity provide a buffer not only against environmental variability as such, but also against variability in patterns of environmental variability. This is precisely what various forms of niche construction do. We mentioned earlier how relocation and migration may buffer against unfavourable patterns of variability; however, not all responses to suboptimal variability need involve seeking out different spatial locations. Some can do so by seeking out new niches in the same spatially located environment.

An increase in 'nestedness' (multiple entities joining to form a greater compositional whole) has been hypothesized to be an example of this, through the mechanism of 'self-extending symbiosis' (Kitano and Oda 2006). Kitano and Oda propose that plastic systems (and specifically, evolvable robust systems) increasingly improve their response to environmental perturbation by integrating foreign biological forms (such as genes and microorganisms). This allows sensitivity to novel environmental variables (or 'inputs' in their terminology).

A lineage thus transformed by self-extending symbiosis would no longer be in direct competition with less nested organisms. Thus, increased nestedness may not only allow for greater viability in suboptimal environments (without guaranteeing viability), but may also allow an organism to avoid direct competition with less nested organisms, thus offering one route to avoid the challenges of unfavourable variability. How large is the probability that a lineage will increase in plasticity in such a way that it is not only responds to variability, but also to variability in patterns of variability? It does not matter how often this occurs: once it occurs, the lineage is robust

against unfavourable environments.

This robustness need not imply that such a lineage is guaranteed to flourish or even survive in all environments. Catastrophic mass-extinctions may wipe out all life, regardless of the degree of plasticity. However, given the multidimensionality and complexity of the environment, all patterns of variability in environmental variables may be expected, and because some forms of plasticity are not only adaptive to specific patterns of variability, but also to variability in patterns of variability, lineages with such forms of plasticity are unlikely to be at such a competitive disadvantage as to become extinct.

Note that in this way it is not necessary for such lineages to be robust against extinction as such; it is only necessary that they are never at a great competitive disadvantage. Of course, in actual evolutionary history, some particularly hostile environments may wipe out lineages with this type of plasticity. In such a case, the trend in plasticity would indeed undergo a local reversal along that area in the space of evolutionary trajectories. However, we need not consider the occurrence of hostile environments that would repeatedly wipe out all lineages with robust plasticity increases, since if no single plastic response is viable, then no fixed response will be viable either.[10]

4.6 Trend vs. Tendency

The combination of increased plasticity during periods of environmental variability and robustness during subsequent suboptimal periods is sufficient to imply a tendency for plasticity to accumulate, without having to make specific assumptions about what types of environment occur. How-

[10]There is some controversy as to whether fixed development may produce extreme phenotypes that would be impossible to produce through plastic phenotypes (DeWitt et al. 1998). However, this is controversial because plasticity likely characterizes all phenotypes, and the effect of canalization might be simply to allow an extreme phenotype to be reliably produced. Further, consider life forms in extreme conditions, for example, extremophile bacteria and other organisms living in and around hydrothermal vents in the ocean. Even such extreme but yet life-supporting conditions are characterized by variability, and we see surprisingly rich ecosystems emerging in these conditions, where species have varying degrees of plasticity. No matter how extreme, as long as an environment is life-supporting, it is characterized by variability, and, given the action of natural selection alone, we may expect increases in plasticity.

ever, a causal tendency does not automatically translate into an *actual trend*, and many changes in the environment could disrupt a trend or influence the shape it takes. An analysis of the nature of the environment and of natural selection can only establish a causal tendency or an idealized, *expected trend*; what the actual trend will look like is a different matter. Even though an investigation of actual trends in plasticity is not the purpose here, it may be helpful to mention some factors that go into determining actual trends other than the tendency outlined above.

The *actual* conditions for a plasticity increase — such as the probability of optimal patterns of variability, or developmental constraints — may be such that the increase can only happen in a limited number of lineages at a time, at irregular intervals, and that the trend may undergo interruptions and reversals.

In general, there could be all sorts of costs that would shape a trend in plasticity (see DeWitt et al. 1998). Some costs are taken to be caused by the development and maintenance of the structures needed for plasticity ('constitutive' costs). Other costs are 'induced' by switching phenotype (for example, switching breeding date for birds, or sex changes for fish). Regardless of the mechanism underlying the costs, the 'jack-of-all-trades' principle is often a motivation for introducing costs: the principle that the optimally adaptive phenotype should perform better when expressed by a specialized, non-plastic individual than when expressed by a plastic individual.

Whether such costs are in fact widespread is controversial and the subject of active research. In principle, if plasticity is the default state – and if non-plasticity requires an expensive suppression of variation through canalization – this removes one obstacle from generalizing from the idealized tendency argued for here to empirical reality. An interesting meta-analysis conducted by van Buskirk and Steiner (Van Buskirk and Steiner 2009) found no widespread costs associated with plasticity. This result is discussed in Murren et al. 2015, who argue that the costs of plasticity are negligible for the majority of organisms, becoming considerable only for organisms with larger brains or complex immune systems. Instead, they found the main limiting factor in the evolution of plasticity to be the presence or absence of optimal selective conditions.

4.7 Comparison with Trends in General Adaptations

The chaotic and multidimensional character of variability of the environment means that patterns of variability are variable as well, over space and over time. Even when described by physical variables alone, the environment is too complex and chaotic to be characterized by a single pattern of variability; this is even more the case when biotic interactions are added into the mix. This 'higher-level' variability is simply another form of environmental variability, and certain forms of plasticity do not provide a buffer either against environmental variability, or against variability in the patterns of environmental variability.

This increase in the probability of viability is lacking in general adaptations, such as size. A large proportion of environments may be resource-rich, or competitive; however, in resource-poor environments, size is not only not optimally adaptive, but it is also not viable. General adaptations produce specific phenotypic responses to common features in environments, but lack the buffering capacity that comes with the ability to hedge over possible phenotypic responses.

Nonetheless, in reality it may be difficult to distinguish trends in general adaptions from trends in plasticity, as the former may be the result of variability selection. For example, Gotanda et al. (2015), in investigating Cope's rule, show that the actual rates of phenotypic change at microevolutionary scales show no bias towards increased body size, despite Cope's rule being reasonably well-confirmed at the macroevolutionary scale. They suggest that the selection for increased body-size is episodic, and that these rare events would often not be captured in microevolutionary studies, due to the short time-scales involved (Gotanda et al. 2015: 1350). Such a pulsed selection pattern (see also Vrba 1993) would be naturally compatible with an intermittently driven variability-selection model, but would be difficult to explain as a result of selection for general adaptations, as the latter does not entail intermittent selection, nor robustness in suboptimal environments.

5 Conclusion

The argument presented is almost wholly abstract, only relying on the existing theoretical understanding of phenotypic plasticity and on facts

about variability in the environment. First, variability in life-supporting environments is ubiquitous and consists of many different temporal and spatial patterns of variability in the various environmental variables.

Second, plasticity is an adaptation that takes advantage of this fact, in a great variety of ways, corresponding to different patterns of variability in different environmental variables. The resulting tendency for plastic responses to accumulate is a reflection of the fact that, in an environment that is variable across many degrees of freedom, a selective advantage can be obtained by increasing sensitivity to information present in an environment and by responding flexibly to it. Such a tendency would correspond to a trend over a succession of lineages (possibly marked by branching events), where plastic responses to various environmental variables are accumulated.

Environments may conspire to interrupt and break such a trend. The actual trend is not independent of what environments happen to transpire. An actual trend in plasticity may be halted or even reversed among various lineages, and may only continue in a fraction of lineages where novel plastic responses to variability are also robust against variability in patterns of variability. However, the succession of environment-types is insufficient to guarantee that a trend in plasticity can be expected to break at some point, in contrast to trends in general adaptations. In an idealized setting, where among other conditions all patterns of variability are equiprobable, a trend in plasticity would be expected to be robust against suboptimal environments, as changes in patterns of environmental variability are simply a different type of environmental variability, for which plastic responses exist. Put differently, given the nature of natural selection and the nature of environmental structure alone, there is no impediment to a trend in plasticity, while there is such an impediment to a trend in a general adaptation.

Conclusion

We frequently engage in speculation about life on other planets, and what life may look many years from now. All such speculation is grounded in a certain understanding of the shape of evolutionary history; yet it is not immediately clear if anything definite can be said about this shape.

There is a proliferation of different interpretations of evolutionary history – different reconstructions of life's plot, if you will – and some of these interpretations reach directly opposed conclusions, despite being based on the same empirical facts and the same biological theories.

Philosophers have not yet actively pursued fundamental conceptual issues facing interpretations (or 'views', or 'perspectives') of evolutionary history. The primary objective of this dissertation was not to defend this or that account, but rather to show a method by which these fundamental issues can be identified and analyzed constructively.

With a hard-nosed skeptic in mind who would be wary of lofty generalizations about complex phenomena we can never observe, I identified two main conceptual issues. The first was the fact that claims about the contingency of evolutionary outcomes depend on how these outcomes and the evolutionary process itself is described. However, instead of consigning ourselves to the conclusion that nothing valuable is to be said beyond 'it depends', we set out to map the different ways in which the contingency of outcomes changes as the phenomena are described in more and in less detail, and as broader or narrower subsets of evolutionary history are taken into account.

That such an analysis is useful to pursue, I attempted to show by applying it to two of the most prominent intepretations of evolutionary history, those of Gould and Conway Morris. According to how their claims about evolutionary history are analyzed, one can arrive at different conclusions about the contingency of evolutionary outcomes.

The second issue concerned whether lofty generalizations about evolutionary history could be viewed as legitimate or not. Evolutionary history is a complex mess of unrelated causal processes, and for every generalization there is an exception. Yet, here again, I sought ways to resist the skeptical conclusion that, therefore, it is a hopeless enterprise to search for generalizations over evolutionary history.

Limiting the scope of the investigation to the ways in which environmental complexity can interrupt trends in evolution by natural selection, I considered first whether and how natural selection may be expected to give rise to trends at all. Some philosophers have rejected that natural selection is a cause at all, and that all evolution is simply an accumulation of births and deaths. If any trend occurs at all, it is due to a confluence of unrelated causal processes that could easily not have occurred. I argued against this by showing (basically) that these philosophers overlooked the issue of time-scale: causal processes may make a difference for reproductive outcome at a time-scale of a single generation without them making a difference at the time-scale of mulitple generations. With this distinction in mind, one can argue that natural selection causes a population to tend towards equilibrium.

If a yet longer time-scale is taken – not that of multiple generations in a single environment, but that of many species and genuses across many environments – the challenge of causal complexity becomes much more difficult to overcome. Drawing on the phenomenon of phenotypic plasticity and niche construction, I argued that selection for plasticity 'feeds' on this complexity and variability in the environment. The trend in plasticity is unique in this regard, as complexity and variability serve to interrupt trends in other types of adaptation.

What the latter argument would mean beyond an idealized evolution by natural selection, and for evolutionary history as a whole is a question beyond the scope of this dissertation. The trend in plasticity has a certain natural compatibility with the fundamental concepts of natural selection and environmental structure, and if anything, would serve to define certain parameters within which large-scale trends in evolutionary history may be sought. In this way, this may orient future research into the twists and turns of life's plot.

Bibliography

Aboulafia, M. (1980). Engels, Darwin, and Hegel's idea of contingency. *Studies in East European Thought 21*(3), 211–219.

Alroy, J. (1998). Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science 280*(5364), 731–734.

Antonovics, J. and P. H. van Tienderen (1991). Ontoecogenophyloconstraints? The chaos of constraint terminology. *Trends in Ecology and Evolution 6*(5), 166 – 168.

Aristotle (1997). *Aristotle's Poetics: Translated and with a Commentary by George Whalley*. McGill-Queen's Press-MQUP.

Arnold, A. J., D. C. Kelly, and W. C. Parker (1995). Causality and Cope's rule: Evidence from the planktonic foraminifera. *Journal of Paleontology 69*(2), 203–210.

Arnold, S. J. (1992). Constraints on phenotypic evolution. *The American Naturalist 140*, S85–S107.

Arthur, B. W. (1994). *Increasing Returns and Path Dependence in the Economy*. University of Michigan Press.

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences 21*(4), 547–569.

Ayala, F. J. (1974). The concept of biological progress. In F. J. Ayala and T. G. Dobzhansky (Eds.), *Studies in the Philosophy of Biology: Reduction and Related Problems*, pp. 339. University of California Press.

Ayala, F. J. (1988). Can "progress" be defined as a biological concept? In M. H. Nitecki (Ed.), *Evolutionary Progress?*, pp. 75–96. The University of Chicago Press.

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist 30*(354), 441–451.

Barros, D. B. (2008). Natural selection as a mechanism. *Philosophy of Science 75*(3), 306–322.

Bassanini, A. and G. Dosi (1999). When and how chance and human will can twist the arms of Clio. Lem Papers Series, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Baumgartner, M. (2010). Interventionism and epiphenomenalism. *Canadian Journal of Philosophy 40*(3), 359–383.

Beatty, J. (1995). The evolutionary contingency thesis. In G. Wolters and J. G. Lennox (Eds.), *Concepts, Theories, and Rationality in the Biological Sciences*, pp. 45–81. University of Pittsburgh Press.

Beatty, J. (2006). Replaying life's tape. *Journal of Philosophy 103*(7), 336–362.

Beatty, J. and E. C. Desjardins (2009). Natural selection and history. *Biology and Philosophy 24*(2), 231–246.

Ben-Menahem, Y. (1997). Historical contingency. *Ratio 10*(2), 99–107.

Bergson, H. (1911). *Creative Evolution*. Henry Holt and Company.

Bonner, J. (1988). *The Evolution of Complexity by Means of Natural Selection*. Princeton University Press.

Bouchard, F. and A. Rosenberg (2004). Fitness, probability and the principles of natural selection. *British Journal for the Philosophy of Science 55*(4), 693–712.

Brading, K. and E. Castellani (2007). Symmetries and invariances in classical physics. In J. Butterfield and J. Earman (Eds.), *Handbook of the Philosophy of Science. Philosophy of Physics*, pp. 1331–1367. North Holland, Elsevier.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in genetics 13*, 115–155.

Brandon, R. N. (1990). *Adaptation and Environment*. Princeton University Press.

Brandon, R. N. (2005). The difference between selection and drift: A reply to Millstein. *Biology and Philosophy 20*(1), 153–170.

Brandon, R. N. (2006). The principle of drift: Biology's first law. *Journal of Philosophy 103*(7), 319–335.

Brooks, D. and E. Wiley (1988). *Evolution as Entropy*. Science and its Conceptual Foundations Series. University of Chicago Press.

Brown, R. L. (2014). What evolvability really is. *British Journal for the Philosophy of Science 65*(3), 549–572.

Canfield, M., E. Greene, D. Whitman, T. Ananthakrishnan, et al. (2009). Phenotypic plasticity and the semantics of polyphenism: a historical review and current perspectives. *Phenotypic plasticity of insects: mechanisms and consequences*, 65–80.

Carlson, S. M., E. M. Olsen, and L. A. Vøllestad (2008). Seasonal mortality and the effect of body size: a review and an empirical test using individual data on brown trout. *Functional Ecology 22*(4), 663–673.

Carroll, R. L. (2000). Towards a new evolutionary synthesis. *Trends in Ecology & Evolution 15*(1), 27–32.

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press.

Cartwright, N. (1995). Ceteris paribus laws and socio-economic machines. *The Monist 78*(3), 276–294.

Chaloner, W. and A. Sheerin (1979). Devonian macrofloras. In M. R. House, C. T. Scrutton, and M. G. Bassett (Eds.), *The Devonian System*, pp. 145–161. Paleontological Society Special Paper in Paleontology.

Chevin, L.-M. and G. M. G. M. Lande (2010). Adaptation, plasticity, and extinction in changing environment: Towards a predictive theory. *PLoS Biology 8*(4).

Chown, S. L. and K. J. Gaston (2010). Body size variation in insects: A macroecological perspective. *Biological Reviews 85*(1), 139–169.

Cover, T. M. and J. A. Thomas (2006). *Elements of Information Theory* (2 ed.). John Wiley & Sons.

Crick, F. H. (1968). The origin of the genetic code. *Journal of Molecular Biology 38*(3), 367–379.

Currie, G. (2010). *Narratives and narrators: A philosophy of stories*. Oxford University Press.

Danto, A. (1965). *Analytical Philosophy of History*. Cambridge University Press. Pagination refers to 3rd edition, published as *Narration in Knowledge* in 2007 by Columbia University Press.

Darwin, C. (1868). *The Variation of Animals and Plants under Domestication*, Volume 2. O. Judd.

David, J. R., P. Gibert, and B. Moreteau (2004). Evolution of reaction norms. In T. J. DeWitt and S. M. Scheiner (Eds.), *Phenotypic Plasticity: Functional and Conceptual Approaches*, pp. 50–63. Oxford University Press.

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review 75*(2), 332–337.

Dawkins, R. (1982). *The Extended Phenotype: The Gene as the Unit of Selection*. Oxford University Press.

Dawkins, R. (1986). *The Blind Watchmaker*. Norton.

Dawkins, R. (2004). *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*. A Mariner Book. Houghton Mifflin.

Dennett, D. (1996). *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. A Touchstone Book. Simon & Schuster.

Desjardins, E. (2011a). Historicity and experimental evolution. *Biology and Philosophy 26*(3), 339–364.

Desjardins, E. (2011b). Reflections on path dependence and irreversibility: Lessons from evolutionary biology. *Philosophy of Science 78*(5), 724–738.

Desjardins, E. (2015). Historicity and ecological restoration. *Biology and Philosophy 30*(1), 77–98.

Desmond, H. (2016). Symmetry breaking and the emergence of path-dependence. *Synthese*, 1–31.

DeWitt, T. J. and S. M. Scheiner (2004). Phenotypic variation from single genotypes: A primer. In T. J. DeWitt and S. M. Scheiner (Eds.), *Phenotypic Plasticity: Functional and Conceptual Approaches*, pp. 1–9. Oxford University Press.

DeWitt, T. J., A. Sih, and D. S. Wilson (1998). Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution 13*(2), 77 – 81.

Dibattista, J. D., K. A. Feldheim, S. H. Gruber, and A. P. Hendry (2007). When bigger is not better: Selection against large size, high condition and fast growth in juvenile lemon sharks. *Journal of Evolutionary Biology 20*(1), 201–212.

Dickins, T. E. and R. A. Barton (2013). Reciprocal causation and the proximate–ultimate distinction. *Biology and Philosophy 28*(5), 747–756.

Dollo, L. (1905). Les dinosauriens adaptés à la vie quadrupède secondaire. *Mémoires de la Société du Géologie 19*, 441–448.

Dombroskie, S. L. and L. W. Aarssen (2010). Within-genus size distributions in angiosperms: Small is better. *Perspectives in Plant Ecology, Evolution and Systematics 12*(4), 283 – 293.

Dowe, P. (2000). *Physical Causation*. Cambridge University Press.

Eimer, G. H. T. (1898). *On orthogenesis and the impotence of natural selection in species formation*. Chicago: Open Court.

Elgin, M. (2006). There may be strict empirical laws in biology, after all. *Biology and Philosophy 21*(1), 119–134.

Elliott, S. (1989). Is the theory of natural selection unprincipled? A reply to Shimony. *Biology and Philosophy 4*(3), 275–279.

Ellis, B. J., A. J. Figueredo, B. H. Brumbach, and G. L. Schlomer (2009). Fundamental dimensions of environmental risk. *Human Nature 20*(2), 204–268.

Endler, J. (1986). *Natural Selection in the Wild*. Monographs in population biology. Princeton University Press.

England, J. L. (2013). Statistical physics of self-replication. *The Journal of Chemical Physics 139*(12).

Ereshefsky, M. (1992). The historical nature of evolutionary theory. In M. H. Nitecki and D. V. Nitecki (Eds.), *History and Evolution*, pp. 81–99. State University of New York Press.

Ereshefsky, M. (2012). Homology thinking. *Biology and Philosophy 27*(3), 381–400.

Falk, D. (2011). *The Fossil Chronicles: How Two Controversial Discoveries Changed Our View of Human Evolution*. University of California Press.

Fisher, Ronald A., S. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh 52*, 399–433.

Fodor, J. A. (1991). You can fool some of the people all of the time, everything else being equal: Hedged laws and psychological explanation. *Mind 100*(397), 19–34.

Forber, P. and K. Reisman (2007). Can there be stochastic evolutionary causes? *Philosophy of Science 74*(5), 616–627.

Forsman, A. (2015). Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity 115*(4), 276–284.

Freund, P. G. O. (2010, August). Emergent Gauge Fields. *arXiv*, 1008.4147.

Futuyma, D. J. (1998). *Evolutionary Biology* (2nd ed.). Sinauer Associates.

Gardiner, P. L. (1952). *The Nature of Historical Explanation*. Greenwood Press.

Gavrilets, S. (2004). *Fitness Landscapes and the Origin of Species*. Princeton University Press.

Geison, G. L. (1969). Darwin and heredity: the evolution of his hypothesis of pangenesis. *Journal of the History of Medicine and Allied Sciences XXIV*(4), 375–411.

Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology 42*(4), 515–553.

Gilbert, S. F. (2000). *Developmental Biology* (6th ed.). Developmental Biology. Palgrave Macmillan.

Gildenhuys, P. (2014). Arbitrariness and causation in classical population genetics. *British Journal for the Philosophy of Science 65*(3), 429–444.

Glennan, S. (2009). Productivity, relevance and natural selection. *Biology and Philosophy 24*(3), 325–339.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.

Gotanda, K. M., C. Correa, M. M. Turcotte, G. Rolshausen, and A. P. Hendry (2015). Linking macrotrends and microrates: Re-evaluating microevolutionary support for Cope's rule. *Evolution 69*(5), 1345–1354.

Gould, S. (1996). *Full House*. Harmony Books.

Gould, S. J. (1989). *Wonderful Life: The Burgess Shale and the Nature of History*. W. W. Norton & Co.

Grove, M. (2011). Change and variability in plio-pleistocene climates: modelling the hominin response. *Journal of Archaeological Science 38*(11), 3038 – 3047.

Grove, M. (2014, August). Evolution and dispersal under climatic instability: A simple evolutionary algorithm. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems 22*(4), 235–254.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese 156*(3), 563–585.

Hansen, T. F. and S. H. Orzack (2005). Assessing current adaptation and phylogenetic inertia as explanations of trait evolution: The need for controlled comparisons. *Evolution 59*(10), 2063–2072.

Hegel, G. W. F. (1996). *Hegel's Lectures on the History of Philosophy*. Humanities Press International.

Hempel, C. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press.

Hempel, C. G. (1942). The function of general laws in history. *Journal of Philosophy 39*(2), 35–48.

Hempel, C. G. and P. Oppenheim (1948). Studies in the logic of explanation. *Philosophy of Science 15*(2), 135–175.

Hull, D. L. (1986). On human nature. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 3–13.

Hull, D. L. (1997). The ascent of ideas in evolution. review of *Monad to Man: The Concept of Progress in Evolutionary Biology* by Michael Ruse. *Nature 385*(6616), 497—498.

Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson and R. A. Hinde (Eds.), *Growing Points in Ethology*, pp. 303–317. Cambridge University Press.

Huneman, P. (2010). Topological explanations and robustness in biological sciences. *Synthese 177*(2), 213–245.

Huneman, P. (2012). Natural selection: A case for the counterfactual approach. *Erkenntnis 76*(2), 171–194.

Huxley, J. (1942). *Evolution: The Modern Synthesis*. John Wiley & Sons.

Irwin, L. N. and D. Schulze-Makuch (2010). *Cosmic Biology: How Life Could Evolve on Other Worlds*. Springer Science and Business Media.

Jablonski, D. (1996). Body size and macroevolution. In D. Jablonski, D. H. Erwin, and J. H. Lipps (Eds.), *Evolutionary Paleobiology*, pp. 256–289. University of Chicago Press.

Jenkin, F. (1867). The Origin of Species. Volume 46, pp. 277–318.

Johnson, K., F. McKinney, and M. Sorenson (1999). Phylogenetic constraint on male parental care in the dabbling ducks. *Proceedings of the Royal Society B: Biological Sciences 266*(1421), 759–763.

Kaplan, J. (2008). The end of the adaptive landscape metaphor? *Biology and Philosophy 23*(5), 625–638.

Kaplan, J. M. (2013). 'Relevant similarity' and the causes of biological evolution: Selection, fitness, and statistically abstractive explanations. *Biology and Philosophy 28*(3), 405–421.

Kauffman, S. A. (1993). *The Origins of Order Self-Organization and Selection in Evolution*. Oxford University Press.

Kearney, M., R. Shine, and W. P. Porter (2009). The potential for behavioral thermoregulation to buffer "cold-blooded" animals against climate warming. *Proceedings in the National Academy of Sciences 106*(10), 3835–3840.

Kettlewell, H. B. D. (1955). Further selection experiments on industrial melanism in the Lepidoptera. *Heredity 10*, 287–301.

Kim, J. (2005). Laws, causation, and explanation in the special sciences. *History and Philosophy of the Life Sciences 27*(3/4), 325–338.

Kim, J. (2012). Against laws in the special sciences. *Journal of Philosophical Research 37*(Supplement), 103–122.

Kitano, H. and K. Oda (2006). Self-extending symbiosis: A mechanism for increasing robustness through evolution. *Biological Theory 1*(1), 61–66.

LaBarbera, M. (1986). The evolution and ecology of body size. In D. M. Raup and D. Jablonski (Eds.), *Patterns and Processes in the History of Life: Report of the Dahlem Workshop on Patterns and Processes in the History of Life*, pp. 69–98. Berlin, Heidelberg: Springer Berlin Heidelberg.

Laland, K., T. Uller, M. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka, J. Odling-Smee, G. A. Wray, and H. E. e. a. Hoekstra (2014). Does evolutionary theory need a rethink? *Nature 514*, 161–164.

Laland, K. N., F. J. Odling-Smee, and M. W. Feldman (1996). The evolutionary consequences of niche construction: A theoretical investigation using two-locus theory. *Journal of Evolutionary Biology 9*(3), 293–316.

Laland, K. N., J. Odling-Smee, W. Hoppitt, and T. Uller (2013). More on how and why: cause and effect in biology revisited. *Biology and Philosophy 28*(5), 719–745.

Laland, K. N. and M. J. O'Brien (2010). Niche construction theory and archaeology. *Journal of Archaeological Method and Theory 17*(4), 303–322.

Laland, K. N. and K. Sterelny (2006). Perspective: seven reasons (not) to neglect niche construction. *Evolution 60*(9), 1751–1762.

Lamsdell, J. C. and S. J. Braddy (2010). Cope's rule and Romer's theory: patterns of diversity and gigantism in eurypterids and palaeozoic vertebrates. *Biology Letters 6*(2), 265–269.

Lange, M. (1993). Natural laws and the problem of provisos. *Erkenntnis 38*(2), 233–248.

Lankford, T. E., J. M. Billerbeck, and D. O. Conover (2001). Evolution of intrinsic growth and energy acquisition rates: Trade-offs with vulnerability to predation in menidia. *Evolution 55*(9), 1873–1881.

Lenski, R. E. (2004). The future of evolutionary biology. *Ludus Vitalis 12*, 67–89.

Lewis, D. (1986). Postscripts to 'causation'. In D. Lewis (Ed.), *Philosophical Papers Vol. II*. Oxford University Press.

Lewontin, R. (1983). Gene, organism, and environment. In D. S. Bendall (Ed.), *Evolution from Molecules to Men*, pp. 273–286. Cambridge University Press.

Lewontin, R. C. (2000). *The Triple Helix: Gene, Organism, and Environment*. Harvard University Press.

Louis, A. A. (2016). Contingency, convergence and hyper-astronomical numbers in biological evolution. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 58*, 107 – 116. Special Issue: *Replaying the Tape of Life: Evolution and Historical Explanation*.

Love, A. C. (2003). Evolutionary morphology, innovation, and the synthesis of evolutionary and developmental biology. *Biology and Philosophy 18*(2), 309–345.

Love, A. C. (2015). Evolutionary developmental biology: Philosophical issues. In T. Heams, P. Huneman, L. Lecointre, and M. Silberstein (Eds.), *Handbook of Evolutionary Thinking in the Sciences*, pp. 265–283. Springer.

Matthen, M. (2009). Drift and statistically abstractive explanation. *Philosophy of Science 76*(4), 464–487.

Matthen, M. and A. Ariew (2002). Two ways of thinking about fitness and natural selection. *Journal of Philosophy 99*(2), 55–83.

Matthen, M. and A. Ariew (2009). Selection and causation. *Philosophy of Science 76*(2), 201–224.

McKitrick, M. C. (1993). Phylogenetic constraint in evolutionary theory: Has it any explanatory power? *Annual Review of Ecology and Systematics 24*, 307–330.

McShea, D. W. (1994). Mechanisms of large-scale evolutionary trends. *Evolution 48*(6), 1747–1763.

McShea, D. W. (1996). Perspective: Metazoan complexity and evolution: Is there a trend? *Evolution 50*(2), 477–492.

McShea, D. W. (2000). Functional complexity in organisms: parts as proxies. *Biology and Philosophy 15*(5), 641–668.

McShea, D. W. (2005). The evolution of complexity without natural selection, a possible large-scale trend of the fourth kind. *Paleobiology 31*(2), 146–156.

McShea, D. W. and R. N. Brandon (2010). *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems*. University of Chicago Press.

Millstein, R. L. (2006). Natural selection as a population-level causal process. *British Journal for the Philosophy of Science 57*(4), 627–653.

Mishler, B. D. and R. N. Brandon (1987). Individuality, pluralism, and the phylogenetic species concept. *Biology and Philosophy 2*(4), 397–414.

Moran, N. A. (1992). The evolutionary maintenance of alternative phenotypes. *American Naturalist*, 971–989.

Moret, B. M. E., L. Nakhleh, T. Warnow, C. R. Linder, A. Tholse, A. Padolina, J. Sun, and R. Timme (2004, January). Phylogenetic networks: Modeling, reconstructibility, and accuracy. *IEEE/ACM Trans. Comput. Biol. Bioinformatics 1*(1), 13–23.

Morris, S. (1998). *The Crucible of Creation: The Burgess Shale and the Rise of Animals*. Oxford University Press.

Morris, S. (2003). *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge University Press.

Müller, G. B. (2007). Evo–devo: extending the evolutionary synthesis. *Nature Reviews Genetics 8*(12), 943–949.

Murren, C. J., J. R. Auld, H. Callahan, C. K. Ghalambor, C. A. Handelsman, M. A. Heskel, J. G. Kingsolver, H. J. Maclean, J. Masel, and H. e. a. Maughan (2015). Constraints on the evolution of phenotypic plasticity: Limits and costs of phenotype and plasticity. *Heredity 115*(4), 293–301.

Nicoglou, A. (2015). The evolution of phenotypic plasticity: Genealogy of a debate in genetics. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 50*, 67–76.

Odling-Smee, F., K. Laland, and M. Feldman (2003). *Niche Construction: The Neglected Process in Evolution*. Monographs in population biology. Princeton University Press.

Payne, J. L., A. G. Boyer, J. H. Brown, S. Finnegan, M. Kowalewski, R. A. Krause, S. K. Lyons, C. R. McClain, D. W. McShea, P. M. Novack-Gottshall, et al. (2009). Two-phase increase in the maximum size of life over 3.5 billion years reflects biological innovation and environmental opportunity. *Proceedings of the National Academy of Sciences 106*(1), 24–27.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Pence, C. H. and G. Ramsey (2013). A new foundation for the propensity interpretation of fitness. *British Journal for the Philosophy of Science 64*(4), 851–881.

Pence, C. H. and G. Ramsey (2015). Is organismic fitness at the basis of evolutionary theory? *Philosophy of Science 82*(5), 1081–1091.

Pierson, P. (2004). *Politics in Time: History, Institutions, and Social Analysis*. Princeton University Press.

Pigliucci, M. (2001). *Phenotypic Plasticity: Beyond Nature and Nurture*. Johns Hopkins University Press.

Pigliucci, M. (2007). Do we need an extended evolutionary synthesis? *Evolution 61*(12), 2743–2749.

Pigliucci, M. and J. Kaplan (2006). *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Theory*. University of Chicago Press.

Pigliucci, M. and G. Muller (2010). *Evolution – the Extended Synthesis*. MIT Press.

Plutynski, A. (2008). The rise and fall of the adaptive landscape? *Biology and Philosophy 23*(5), 605–623.

Potts, R. (1998). Environmental hypotheses of hominin evolution. *Yearbook of Physical Anthropology 41*, 93–136.

Powell, R. (2009). Contingency and convergence in macroevolution: A reply to John Beatty. *Journal of Philosophy 106*(7), 390–403.

Powell, R. (2012). Convergent evolution and the limits of natural selection. *European Journal for Philosophy of Science 2*(3), 355–373.

Price, T. D., A. Qvarnström, and D. E. Irwin (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B: Biological Sciences 270*(1523), 1433–1440.

Prothero, D. R. (1997). Inevitable humans? or hidden agendas? Review of *Life's Solution: Inevitable Humans in a Lonely Universe* by Simon Conway Morris. *Skeptic 10*(3), 54—57.

Ramsey, G. (2015). The causal structure of evolutionary theory. *Australasian Journal of Philosophy 94*(3), 421–434.

Ramsey, G. and A. S. Peterson (2012). Sameness in biology. *Philosophy of Science 79*(2), 255–275.

Reisman, K. and P. Forber (2005). Manipulation and the causes of evolution. *Philosophy of Science 72*(5), 1113–1123.

Rosslenbroich, B. (2006). The notion of progress in evolutionary biology – the unresolved problem and an empirical suggestion. *Biology and Philosophy 21*(1), 41–70.

Ruse, M. (1990). Are pictures really necessary? The case of Sewall Wright's "adaptive landscape". *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990*, 63–77.

Ruse, M. (1996). *Monad to Man: The Concept of Progress in Evolutionary Biology*. Harvard University Press.

Russell, B. (1912). On the notion of cause. *Proceedings of the Aristotelian Society 7*, 1–26.

Salisbury, F. B. (1969). Natural selection and the complexity of the gene. *Nature 224*, 342–343.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

Salmon, W. C. (1971). *Statistical Explanation and Statistical Relevance*. University of Pittsburgh Press.

Schank, J. C. and W. C. Wimsatt (1986). Generative entrenchment and evolution. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1986*, 33–60.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics 24*(1), 35–68.

Scheiner, S. M. and R. F. Lyman (1989). The genetics of phenotypic plasticity: Heritability. *Journal of Evolutionary Biology 2*, 95 – 107.

Schlichting, C. and M. Pigliucci (1998). *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer.

Schulze-Makuch, D., A. Schulze-Makuch, and J. M. Houtkooper (2015). The physical, chemical and physiological limits of life. *Life 5*(3), 1472—1486.

Scriven, M. (1959). Truisms as the grounds for historical explanations. In P. L. Gardiner (Ed.), *Theories of History*, pp. 443–475. Free Press.

Scriven, M. (1962). Explanations, predictions and laws. In H. Feigl and G. Maxwell (Eds.), *Scientific Explanation, Space, and Time*, pp. 75–96. The University of Chicago Press.

Shanahan, T. (2011). Phylogenetic inertia and Darwin's higher law. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 42*(1), 60–68.

Shapiro, L. A. and E. Sober (2007). Epiphenomenalism: The do's and don'ts. In G. Wolters and P. K. Machame (Eds.), *Thinking about Causes: From Greek Philosophy to Modern Physics*, pp. 235–264. University of Pittsburgh Press.

Siepielski, A. M., J. D. DiBattista, and S. M. Carlson (2009). It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters 12*(11), 1261–1276.

Simpson, G. G. (1944a). *The Meaning of Evolution*. Yale University Press. (Revised Edition 1971).

Simpson, G. G. (1944b). *Tempo and Mode in Evolution*. Columbia University Press.

Skipper, R. A. (2004). The heuristic role of Sewall Wright's 1932 adaptive landscape diagram. *Philosophy of Science 71*(5), 1176–1188.

Skipper, R. A. and R. L. Millstein (2005). Thinking about evolutionary mechanisms: Natural selection. *Studies in History and Philosophy of Biological and Biomedical Sciences 36*(2), 327–347.

Smart, J. J. C. (1963). *Philosophy And Scientific Realism*. Humanities Press.

Smith, D. L. (2011). *Less Than Human: Why We Demean, Enslave, and Exterminate Others*. St. Martin's Press.

Smith, J. M. (1970). Natural selection and the concept of a protein space. *Nature 225*, 563–564.

Smith, J. M., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert (1985). Developmental constraints and evolution. *Quarterly Review of Biology 60*(3), 265–287.

Smith, L. D. and A. R. Palmer (1994). Effects of manipulated diet on size and performance of brachyuran crab claws. *Science 264*(5159), 710–712.

Sober, E. (1983). Equilibrium explanation. *Philosophical Studies 43*(2), 201–210.

Sober, E. (1984). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. University of Chicago Press.

Sober, E. (1988). *Reconstructing The Past: Parsimony, Evolution, and Inference*. MIT Press.

Sober, E. (1994). The adaptive advantage of learning and a priori prejudice. In E. Sober (Ed.), *From a Biological Point of View*, pp. 50–69. Cambridge University Press.

Sober, E. (2013). Trait fitness is not a propensity, but fitness variation is. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 44*(3), 336–341.

Sober, E. and M. Steel (2011). Entropy increase and information loss in Markov models of evolution. *Biology and Philosophy 26*(2), 223–250.

Spirtes, P., C. Glymour, and R. Scheines (2000). *Causation, Prediction, and Search*. MIT Press: Cambridge.

Stephens, C. (2004). Selection, drift, and the 'forces'of evolution. *Philosophy of Science 71*(4), 550–570.

Stephens, C. (2010). Forces and causes in evolutionary theory. *Philosophy of Science 77*(5), 716–727.

Sterelny, K. (2001). *Dawkins vs. Gould: Survival of the Fittest*. Icon Books.

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies 127*(1), 109–166.

Strevens, M. (2003). *Bigger Than Chaos: Understanding Complexity Through Probability*. Harvard University Press.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Strevens, M. (2012). Ceteris paribus hedges: Causal voodoo that works. *Journal of Philosophy 109*(11), 652–675.

Strevens, M. (2016). The reference class problem in evolutionary biology: Distinguishing selection from drift. In C. Pence and G. Ramsey (Eds.), *Chance in Evolution*. University of Chicago Press.

Stroock, D. W. (2005). *An Introduction to Markov Processes*. Springer.

Szathmáry, E. (2006). Path dependence and historical contingency in biology. In A. Wimmer and R. Kössler (Eds.), *Understanding Change: Models, Methodologies, and Metaphors*, pp. 140–157. Palgrave Macmillan.

Taylor, T. B., G. Mulley, A. H. Dills, A. S. Alsohim, L. J. McGuffin, D. J. Studholme, M. W. Silby, M. A. Brockhurst, L. J. Johnson, and R. W. Jackson (2015). Evolutionary resurrection of flagellar motility via rewiring of the nitrogen regulation system. *Science 347*(6225), 1014–1017.

Trussell, G. C. and L. D. Smith (2000). Induced defenses in response to an invading crab predator: an explanation of historical and geographic phenotypic change. *Proceedings of the National Academy of Sciences 97*(5), 2123–2127.

Valen, L. V. (1973). A new evolutionary law. *Evolutionary Theory 1*, 1–30.

Van Buskirk, J. and U. Steiner (2009). The fitness costs of developmental canalization and plasticity. *Journal of Evolutionary Biology 22*(4), 852–860.

Velasco, J. D. and E. Sober (2010). Testing for treeness: Lateral gene transfer, phylogenetic inference, and model selection. *Biology and Philosophy 25*(4), 675–687.

Verlinde, E. (2011). On the origin of gravity and the laws of newton. *Journal of High Energy Physics 2011*(4), 29.

Vermeij, G. J. (1987). *Evolution and Escalation: An Ecological History of Life*. Princeton University Press.

Vermeij, G. J. (1999). Inequality and the directionality of history. *The American Naturalist 153*(3), 243–253.

Vrba, E. S. (1993). Turnover-pulses, the red queen, and related topics. *American Journal of Science 293*(A), 418–452.

Waddington, C. H. (1953). Genetic assimilation of an acquired character. *Evolution 7*(2), 118–126.

Walsh, D. M. (2000). Chasing shadows: Natural selection and adaptation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 31*(1), 135–153.

Walsh, D. M. (2007). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science 74*(3), 281–303.

Walsh, D. M. (2010). Not a sure thing: Fitness, probability, and causation. *Philosophy of Science 77*(2), 147–171.

Walsh, D. M., A. Ariew, and T. Lewens (2002). The trials of life: Natural selection and random drift. *Philosophy of Science 69*(3), 452–473.

Wei, Y., X. Wang, J. Liu, I. Nememan, A. H. Singh, H. Weiss, and B. R. Levin (2011). The population dynamics of bacteria in physically structured habitats and the adaptive virtue of random motility. *Proceedings of the National Academy of Sciences 108*(10), 4047–4052.

Wells, D. A. (2015). The extended phenotype(s): a comparison with niche construction theory. *Biology and Philosophy 30*(4), 547–567.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Whitman, D. W. and A. Agrawal (2009). What is phenotypic plasticity and why is it important? In D. W. Whitman and T. N. Ananthakrishnan (Eds.), *Phenotypic Plasticity in Insects: Mechanisms and Consequences*, pp. 1–65. Taylor and Francis Group: Science Publishers.

Wigglesworth, V. (1961). Insect polymorphism: A tentative synthesis. pp. 103–113.

Wilkins, J. F. and P. Godfrey-Smith (2009). Adaptationism and the adaptive landscape. *Biology and Philosophy 24*(2), 199–214.

Wilson, J. (2007). Newtonian forces. *British Journal for the Philosophy of Science 58*(2), 173–205.

Woltereck, R. (1909). Abteilung vii: Zoologie des meeres (exkl. vertebraten). jahresübersicht der literatur für das jahr 1908. *Internationale Revue der gesamten Hydrobiologie und Hydrographie 2*(6), A77–A104.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics 1*, 356–66.

Young, N. M., B. HallgrÍmsson, and C. Janis (2005). Serial homology and the evolution of mammalian limb covariation structure. *Evolution 59*(12), 2691–2704.